Intelligence Testing

in a Toronto Public School

A Thesis by

CECIL C. GOLDRING, M.A.

INTELLIGENCE TESTING
in a
TORONTO PUBLIC SCHOOL

Intelligence Testing
in a
Toronto Public School

By

Cecil C. Goldring, M.A.

Assistant Master, Bolton Avenue School,
Toronto, Ontario.

A thesis submitted in conformity with the requirements for the Degree of Doctor of Pedagogy in the University of Toronto.

CONTENTS

Chapter	Title	Page
	Introduction	11
1.	An Account of the Children Tested and a Comparison of Results with Teachers' Ratings and Examination Records	29
2.	Class-room Grading	47
3.	The Accomplishment Ratio	62
4.	Comparison of Results of Terman, Otis and Trabue Tests	77
5.	The Constancy of the I. Q.	86
6.	Siblings and Twins	104
7.	The Relation of Intelligence to Earning Power	114
8.	The Face as an Index of Intelligence	125
9.	Concluding Notes	136

INTRODUCTION

THE MEANING AND HISTORY OF INTELLIGENCE TESTS

If an interested parent were to question a teacher regarding the intelligence of his child, the teacher would probably reply that he regarded the child as bright, very bright, average, dull, or very dull. If the parent urged the teacher to be more specific as to the degree of brightness or dullness, the teacher would probably find difficulty in giving a satisfactory answer. He might compare him in brightness or dullness with the extremes of his class; or he might make the comparison between the child and other children well known to both parent and teacher. If the question, instead of being one concerning intelligence, were one as to the height or weight of the child, there would be no difficulty in giving an exact answer. By measurement the height or weight could be exactly determined, even in such small units as eighth-inches or ounces. An intelligence test is an attempt to apply exact measurement to mental processes; just as length, area, volume, weight, temperature, etc., are all subject to exact measurement in terms of certain units.

Before discussing the tests in detail, the meaning of intelligence must be explained. In educational psychology, different investigators attach somewhat different meanings to the word. In order to present different points of view on the subject, the *Journal of Educational Psychology* in March, 1921, secured articles from fourteen leading American psychologists on the subject, "What I Conceive Intelligence to Be,

and by What Means It Can Best Be Measured by Group Tests." Short representative answers from this collection of expert opinion will be quoted. The complete answers may be found in the *Journal of Educational Psychology*, Volume XII.

"An individual is intelligent in proportion as he is able to carry on abstract thinking."—*Dr. Terman, Leland Stanford University.*

"We may define intellect in general as the power of good responses from the point of view of truth or fact, and may separate it according as the situation is taken in gross or abstractly and also according as it is experienced directly or thought of. The power of good responses to abstract qualities and relations rather than gross total facts and to ideas rather than direct experiences may be called the more intellectual variety of intellect."—*Dr. Thorndike, Columbia University.*

"An individual possesses intelligence in so far as he has learned, or can learn to adjust himself to his environment."—*The late Dr. S. S. Colvin, of Columbia University.*

"I have always thought of intelligence as the ability of the individual to adapt himself adequately to relatively new situations in life. It seems to include the capacity for getting along well in all sorts of situations. This implies ease and rapidity in making adjustments and, hence, ease in breaking old habits and in forming new ones."—*Dr. R. Pintner, Columbia University.*

"Frankly, then, the concept of general intelligence is for the writer, primarily, simply a working hypothesis which has been very helpful in the first attack upon problems of prognosis. It should continue

as a working hypothesis until other hypotheses are found to work better."—*Dr. Pressey, University of Indiana.*

"Intelligence is indicated by the capacity to appropriate truth and fact as well as by the capacity to discover them."—*Dr. Henmon, University of Wisconsin.*

"Intelligence seems to be a biological mechanism by which the effects of a complexity of stimuli are brought together and given a somewhat unified effect in behavior. It is a mechanism for adjustment and control, and is operated by internal as well as external stimuli. The degree of a person's intelligence increases with his range of receptivity to stimuli and the consistency of his organization of responses to them."—*Dr. Peterson, George Peabody College for Teachers.*

"Intelligence as judged in everyday life contains at least three psychologically differentiable components; the capacity to inhibit an instinctive adjustment, the capacity to redefine the inhibited instinctive adjustment in the light of imaginally experienced trial and error, the volitional capacity to realize the modified instinctive adjustment into overt behavior to the advantage of the individual as a social animal."—*Dr. Thurstone, late of Carnegie Institute of Technology.*

"Intelligence, then, is the capacity to acquire capacity. . . . Intelligence, except for being a growing thing, is fixed partly by heredity and partly by environmental factors acting before the age of five. . . . Intelligence is neither produced, nor appreciably accelerated, by learning."—*Dr. Woodrow, University of Minnesota.*

"It seems to me that whatever definition we may give to intelligence in the abstract, we are justified from an educational point of view in regarding it as ability to learn, and as measured by the extent to which learning has taken place or may take place." —*Dr. B. R. Buckingham, Ohio State University, Bureau of Educational Research.*

Some of the opinions quoted emphasize certain aspects of the general subject of intelligence. One writer views it from the biological point of view, another from the educational, another from the point of view of submitting a careful, exact definition, while still another admits that he cares little for a precise definition, but is more interested in the use which can be made of the concept of intelligence. Fundamentally, however, one can detect a certain uniformity of view. Probably all would agree that intelligence is a general mental ability operating in many different ways; it is more fully manifested in the higher mental processes than in the lower: it is specially evident in dealing with a new or novel situation: for schoolroom purposes the degree of intelligence becomes evident in the capacity of the pupil to learn, and especially to reason. As Ballard, an English psychologist, points out in his book, "Group Tests of Intelligence": "It is not only a popular belief, but a scientific hypothesis which has received no small amount of experimental support, that the nearer we approach reason, the nearer do we get to the very soul and secret of intelligence."

The problem of intelligence may be considered from another point of view. The term has been defined and there are numerous tests designed by psy-

chologists and labelled "intelligence tests." These tests purport to measure some mental factor. What do they measure? Terman states that his revision of the Binet scale "includes tests of time orientation, of three or four kinds of memory, of apperception, of language comprehension, of knowledge about common objects, of free association, of number mastery, of constructive imagination, and of ability to compare concepts, to see contradictions, to combine fragments into a unitary whole, to comprehend abstract terms, and to meet novel situations."

Speaking generally, an intelligence test measures knowledge, but the knowledge tested is limited to that which the subject of ordinary intelligence cannot possibly avoid; it is thrust upon him in the common intercourse of daily life. Tests also attempt to measure such capacities as sensory capacity; capacity for perceptual recognition; quickness, range or flexibility of association; facility in imagination; span or steadiness of attention; quickness or alertness in response. The tests measure a "something" which grows with a child's growth up to a certain age and then appears to stop; it seems to bear, up to about sixteen, a constant relation to chronological age, *i. e.*, to increase with chronological age at a certain rather definite rate; it is independent of special schooling or special training; and it is closely related to success in school and success in life.

Referring to the history of intelligence tests, as given by Wylie, "A Brief History of Mental Tests," *Teachers' College Record*, January, 1922, one finds evidence of attempts made several centuries ago to determine the mental status of persons charged with crime. In Roman law different types of insanity

were designated by such terms as *demens*, *furiosus*, and *fatuus*. Fitzherbert (1470 - 1538) an English legal writer, advocated judging the mentality of one suspected of being insane by the "capacity of the alleged idiot to count twenty pence, or to tell his age, or who were his father and mother." Swinburne (1560 - 1623) suggested examining such a person to ascertain his ability to "measure a yard of cloth or name the days in the week." In the seventeenth century a legal test proposed to determine responsibility for crime was the "level of understanding of a child fourteen years of age."

In considering the historical background of intelligence tests, mention should also be made of early studies in the fields of phrenology and physiognomy. Near the end of the eighteenth century, Gall, the father of phrenology, taught that by measuring the skull and the configurations of its surface, an exact estimate of the mental powers of the individual might be determined. When Gall was still a boy of fourteen, Lavater, a physiognomist, published, in 1772, an essay claiming that the face was the index of the mind; for example, a nose protruding near the root indicated an aggressive disposition, while one that protruded in the middle indicated a propensity to fight for others. About a century later Lombroso, a criminologist, sought to distinguish the criminal type by studying the physical peculiarities and malformations of those who were confirmed evil-doers; he claimed, for example, that a lack of symmetry in the body or in the features, was often an indication of degeneration. Sir Francis Galton, the eminent English scientist, was convinced that some correspondence existed between intelligence and certain bodily traits.

His researches, as reported in "Inquiries into Human Faculty and Its Development" and "Hereditary Genius," included the study of several physical traits; perhaps his most valuable conclusion in this field of study was that there was a relation between intelligence and the character of the finger prints,—a forerunner of our modern system of identification by finger prints. But Prof. Karl Pearson dealt a blow to investigations of this type by publishing in 1906 the result of an elaborate investigation into the relationship between intelligence and the shape and size of the skull. His conclusion was that, if there were any connection at all between them, it was so slight as to be practically valueless for purposes of inference.

Another chapter in the history of intelligence tests has been contributed by those who have attempted to rate the relative intelligence of various races, and to rate men of outstanding talent according to some scientific standards. Among studies of this kind are the following: Galton's "Hereditary Genius" (1869, 1892) : "British Men of Science" (1874) : Cattell's "American Men of Science" (1903) : Ellis' "British Genius" (1904) : and Wood's "Mental and Moral Heredity in Royalty" (1906).

The rise of experimental psychology and the introduction of mathematical methods into psychological research also contributed largely to the evolution of the modern intelligence test. Psychological measurements really began with Weber and Fechner. Weber's law (1834) might be considered an early landmark in this field. This law is stated as follows in Webster's New International Dictionary : "The law or generalization that the least noticeable increase of a stimulus is a constant proportional of the original

stimulus." In 1860 Fechner published his "Psychophysik," which aroused considerable discussion and controversy and stimulated further investigations. Gradually there were evolved quantitative methods somewhat as we use them to-day. The next important contribution, derived from Fechner's work, was Wundt's "Principles of Physiological Psychology," which appeared in 1874, with a second edition in 1880.

Shortly after 1890 scientists in France, Germany, United States and England became interested in constructing mental tests primarily designed to diagnose insanity and mental abnormality. The next step was to formulate tests which would indicate the mental capacity of sane people as well as of the alleged insane. Certain psychologists saw the advantages which might accrue to educational effort if tests were devised to measure the mental capacity of school children. Tests were then devised more from the standpoint of use in the schools and use with children. This very important field of educational psychology then developed rapidly.

The earliest record of work in the field of educational measurement was that of Reverend George Fisher of the Greenwich Hospital School, England, who, in 1864, used a "Scale Book" which contained the numbers assigned to each degree of proficiency in the various subjects of examination. This "Scale Book" included specimens of writing, spelling, mathematics, and a variety of school subjects, the different specimens being rated for excellence on a scale of 1 to 5. In America the real inventor of the comparative test was Dr. J. M. Rice, who, in 1894, began his investigations of spelling ability.

In the period 1890 - 1905 numerous mental tests

were devised, some to measure definite mental capacities and some to measure general intelligence. Many of them have survived and become incorporated in the present intelligence tests. Collectively they form the material from which many of the present standard intelligence tests were evolved. For an account of these early experiments Whipple's "Manual of Mental and Physical Tests" may be consulted. Some of them will be enumerated here to indicate the fields of investigation. They are: Tests of the Range of Visual Attention: Visual Apprehension: Cancellation Test: Description of an Object: Fidelity of Report Test: Uncontrolled Association: Controlled Association: Logical Relations: Mirror Drawing: Rote Memory: Logical Memory: Linguistic Invention: Completion Method: The Interpretation of Fables: Size of Vocabulary: Range of Information.

In 1905, Binet, a Frenchman, published a scale of intelligence tests which was an improvement on any one previously used. The scale consisted of 30 tests, arranged in order of difficulty. Binet had tried the tests upon some 200 normal children of different ages from 3 to 15 years. In the case of any one of the 30 tests, if it were passed by about two-thirds to three-fourths of the normal children of any age, say 7 years, it was considered a test of 7-year intelligence. In this way he standardized his 30 tests in age groups; he could then speak of a child of 9-year-old intelligence, or 10-year-old intelligence, and such an expression would have a definite meaning. This was a great advance and made it possible to characterize the intelligence of a child in a far more definite way than had hitherto been possible.

In 1908 Binet issued a second series of 56 tests.

Goddard first published an account and an English translation in 1910. In 1911 Binet published his latest revision of the scale; the Binet tests were introduced into America by Goddard, and Terman revised them to American needs in 1915. Terman examined some 2,300 subjects, of whom 1,700 were normal children, and then arranged the tests in age groups, extending from 3 years to 18 years (superior adult); there are 90 tests in his complete scale. Terman's notable addition was the use of the intelligence quotient or I. Q. This is the ratio of the mental age to the chronological age. Thus a 10-year-old child should theoretically reach the 10-year-old standard when tested; his I. Q. is then 100. In reality, however, few have I. Q.'s of exactly 100. Most people are mentally slightly advanced or retarded in comparison with their chronological ages. Thus a 10-year-old child might have only the mental development of an average 9-year-old; then his I. Q. is 90. On the other hand, he may be mentally equal to the average 11-year-old child; his I. Q. is then 110. Terman found that about 60 per cent of all school children test between 90 and 110 I. Q.

Terman's scale of tests aroused interest among American educationists. But very soon the United States entered the Great War. This provided an opportunity to put intelligence testing to a practical use. After some preliminary testing, which justified the claims made for them by their advocates, it was decided to test all the men enlisted in the American army with the exception of field and general officers. Between September, 1917, and January, 1919, about 1,727,000 men were tested. To examine this large number, group tests were compiled and the men were

examined in groups of from 75 to 500. About fifty minutes was required for the examination of the entire group. In some 83,500 special cases an additional individual test was given. It was officially proclaimed that the purpose of the army psychological examinations was not only to eliminate the unfit, but also to assess the intelligence of every soldier in the army: to indicate the kind of training each recruit might most profitably undergo: to permit of the units being so organized that men of the same grade of intelligence could serve in the same regiment; and to pick out those men whose general ability marked them for promotion or whose special abilities fitted them for special kinds of military service.

The army testing was a success. It succeeded so well that the American nation became so firmly convinced of the predictive value of intelligence tests that their use has been extended to nearly every department of life, especially to education. Many other group tests based on that used with the army have been devised, some of the well-known tests being: Otis (1918), Pressey-Pressey (1918), Thorndike (1919), Thurstone (1919), Whipple (1919), Dearborn (1920), Haggerty (1920), Monroe and Buckingham (1920), Terman (1920), and Trabue (1920). Another well-known and elaborately designed series is the National Intelligence Tests (1920). This series was prepared under the direction of the National Research Council and is the joint production of Haggerty, Terman, Thorndike, Whipple and Yerkes.

Especially during the last ten years, many tests have been devised to measure the progress of pupils in the recognized school subjects. Standard scores based on grade are available, and thus it is possible to measure

the progress of the individual pupil or of a class as a unit and to compare the scores with the standard attained by other pupils of the same grade. In a few cases standard scores based on age have also been derived. To mention all of them would require much space. A few of such tests are:

Arithmetic—Courtis (1910): Woody-McCall (1919).
Algebra—Hotz (1918): Monroe (1914): Rogers and Rugg (1919).
Language—Starch: Charters.
Reading—Thorndike (1914): Starch (1915).
Spelling—Starch (1915): Ayres (1915).
Handwriting — Ayres (1912): Freeman (1914): Thorndike (1908).
English Composition — Thorndike (1915): Trabue (1917): Hillegas (1912).

A beginning has recently been made at testing certain qualities of importance in everyday conduct. The names of such usually indicate the qualities tested Examples of this type are: An Individual Will Temperament Test by Downey (1921): The Upton-Chassell Scale of Measuring Habits of Good Citizenship (1919): The Voelker Scale to Test Morality and Trustworthiness (1921).

Intelligence tests have a wide use in the United States at present. Only eighteen years have elapsed since Binet devised his first crude series of intelligence tests, only eight since Terman adapted and revised Binet's scale in America. In the United States at present tests are used in schools; for college entrance at Columbia, Brown, Rutgers, and elsewhere; to help in the selection of employees and executives for business

and industry in many business organizations: also in prisons, juvenile courts, reform schools, and institutions for defectives. Terman, writing in 1922, states that " there is reason to believe that at the present time Binet tests are being given in the United States at the rate of a quarter of a million a year." The following is also a quotation from " Intelligence Tests and School Reorganization " by the same authority:

" Probably a million children in the schools of the United States were given a group mental test during the year 1919-20. In 1920-1921 the number was probably not less than two millions. We may expect the number to exceed five millions within a few years."

Another evidence of their increasing use in the United States is the following quotation from a pamphlet issued on March 1, 1923, by a company which specializes in the publishing of intelligence tests:

" During the past year about three million children were supplied by us with test material of various kinds —an increase of nearly 80% over the preceding year."

In England, Mr. Cyril Burt has taken a leading part in the development of intelligence tests. He began his investigations among school children in Oxford about 1905. His method at first was to test a group of children and then have the teacher arrange these children in order of intelligence, this grading being partly the result of the teacher's empirical judgment and partly the result of school examinations. Burt continued his investigations at Liverpool, and, on the basis of the results of his testing, he arranged a large number of tests he had used in a series graded in order of complexity. He later devised a series of Reasoning Tests, consisting of fifty tests in all. It was Galton, another English psychologist previously referred to,

who first suggested that mental traits generally, and intelligence in particular, would be found to follow the law of normal distribution; the greatest contribution of this psychologist, however, was the recognition of individual differences. To Galton, also, we are mainly indebted for the doctrine of correlation, which has proved so valuable in comparing the respective order of the individuals in two lists of measurements. The doctrine of correlation was further developed by Pearson and a formula for calculating it was put forward by him. Spearman later discovered another simpler formula of correlation, which has been improved upon by Pearson.

There has not been as rapid development of intelligence testing in England as in the United States but steady progress is being made. It is said that the London County Council made the first appointment in 1913 of any city in the world of a psychologist as an official on its staff. Mr. Cyril Burt's recent books, "The Distributions and Relations of Educational Abilities" and "Three Memoranda on Mental and Scholastic Tests," based on the testing of nearly 6,000 individuals from different types of institutions and schools, are reported by reviewers to be works of exceptional worth, and it is prophesied that they will do much to promote educational efficiency.

Several group tests of intelligence have been devised in England. One of the best known is the Northumberland Mental Tests: others are the Columbian Mental Tests: the Chelsea Mental Tests: the Crichton Test and the Simplex Group Intelligence Scale. Burt has also devised several group tests. The Bradford Education Committee has during the past few years used mental tests for the selection of scholarship children.

The Northumberland Education Committee has also adopted a similar plan. Another noteworthy use of them is that in the Civil Service competitive examinations for women clerks in October, 1920, a section labelled "Intelligence Tests" appeared for the first time. Each subsequent Civil Service examination has also contained a section with the same heading. Another evidence of the increasing popularity of intelligence tests in England is found in the fact that the Education Research Committee of the London Head Teachers' Association, after experimenting with mental tests, has pronounced them sound and practicable. The psychological staff of University College, London, has also done considerable experimental work with group tests.

The following quotation dealing with the extension of their use is from Ballard's "Group Tests of Intelligence" (1922):

"The use of group tests of intelligence for selective purposes has not been slow in reaching this side of the Atlantic. Education authorities in England have begun to use them for the award to scholarships to secondary schools. They are being extensively used for the same purpose in Germany, especially at Berlin, Hanover and Leipzig. At Hamburg 1,000 children are tested each spring, and the best of them transplanted to higher nurseries of learning.

"The system of selection by mental tests is not confined to schools and colleges: it has invaded municipal and business life. The Civil Service has partly adopted it, and a few large business firms in London and elsewhere seem to have wholly adopted it."

In Canada little systematic use of intelligence tests has yet been made but there is abundant evidence of a

general interest in the subject. In the province of Ontario, during the past few years, there has been a branch of the Department of Education, under the direction of Dr. S. B. Sinclair, which has had charge of psychological work in the schools. The Ontario Government makes free surveys of any school area desiring to form an auxiliary training class for very backward children, having I. Q.'s between 50 and 75. The school staffs are given definite instructions how to select pupils for special examination. The selected pupils are examined by an educational psychologist from the Department who gives the Terman test, sometimes reinforced by a performance test. During the past four years 27 surveys have been conducted outside the city of Toronto and the number of classes in the province for very backward children has increased from 6 to 72,—a twelvefold increase in four years! The Board of Education of the city of Toronto employs a psychiatrist to work in the schools under its control. For several years, tests have also been used in connection with Toronto's Juvenile Court, a total of 688 children and adults having been given mental examinations by the Court psychologist during the year 1922.

One hears occasionally of individual schools and teachers throughout Ontario, and in the western provinces particularly, using intelligence tests and grading the school partly or wholly on the result of these tests, or using them to place individual pupils where they can do more effective work. Queen Alexandra School, Toronto, a school of 1,300 pupils, has classified pupils for the past two and a half years on the combined result of intelligence tests and teachers' judgments. The aim has been to give each pupil a group intelligence test each year: the group tests used

being the National, Otis, Haggerty, and, to a lesser extent, the Terman. In addition, all exceptional children, either very bright or very dull, are given the individual Terman test. At present (February, 1924) the school has nine "dull normal" classes, in which both the method of instruction and the content differ from the work of the regular classes. No special classes exist for the very bright children.

The Strathcona Model School, Hamilton, is carrying out an experiment designed to have pupils move through the school as mental age groups. In that school the content of each subject in each grade has been divided into three parts. The classes are then arranged so as to make it possible for the bright pupils to advance at a more rapid rate than the dullards.

Thus it is evident that Canadians in various places are experimenting with intelligence tests and their use will doubtless become more general as their value in Canada is more fully demonstrated.

CHAPTER I

AN ACCOUNT OF THE CHILDREN TESTED AND A COMPARISON OF RESULTS WITH TEACHERS' RATINGS AND EXAMINATION RECORDS

In preparation for this study, 509 public school pupils of Toronto were given the Stanford Revision of the Binet-Simon Intelligence Test. This is probably the best known individual intelligence test and its results are generally regarded as very accurate. It was prepared by Dr. Terman of Leland Stanford University; it will be frequently referred to in this study as the "Terman test." A complete account of the tests and the method of administering and scoring them may be found in Terman's "Measurement of Intelligence." The instructions in that book were strictly followed, and every precaution was taken to obtain results which would be as reliable as possible.

All the pupils tested attended Bolton Avenue Public School, Toronto. This school is an average one, attended by the children of working people. Few of the parents are wealthy, but few are in abject poverty. Nearly all the parents were born in Canada or in the British Isles. The country of birth of 380 of the 509 children tested was ascertained. In the case of the remaining 129, it can be definitely stated that from 80% to 90% of them were born in Canada, the remainder being natives of the British Isles. The result was as follows:

Country of Birth	No.
Canada	308
England	43
Scotland	14
Ireland	9
United States	5
Macedonia	1
Total	380

Thus 81% were found to be of Canadian birth, and all the others, with one exception, were born in an English-speaking country.

The social standing and financial condition of the parents can be estimated to some extent by a statement of the father's occupation. The occupation of the fathers of 421 of the 509 pupils was ascertained. The fathers of many of the others are dead. Table I gives the number of children based on the fathers' occupations and the average I. Q. of the children for each occupation.

Table I

Analysis of I. Q.'s Based on Fathers' Occupations

Father's Occupation	No. of Children	Children's Average I. Q.
Policeman	3	117
Bricklayer	4	109
Manufacturer	5	107
Carpenter	21	106
Woodworker	11	103
Minister or lawyer	4	102
Foreman or inspector	16	102
Electrician	5	102

Father's Occupation	No. of Children	Children's Average I. Q.
Clerk	21	101
Florist	6	101
Tailor	14	100
Machinist	24	100
Fireman	5	100
Railway employee	13	100
Chauffeur	3	100
Agent	8	99½
Janitor	6	99
Engineer	11	99
Postman	8	99
Watchman	3	99
Shipper	13	98
Baker	4	97
Printer	10	96
Civic employee	15	96
Storekeeper	23	96
Traveller	6	95
Glass cutter	3	94
Market gardener	3	94
Painter	10	92
Tinsmith	11	91
Driver	17	90
Factory worker	37	90
Expressman	11	90
Salesman	3	89
Barber	7	89
Musician	3	87
Laborer	21	84
Motorman or conductor	5	83
Miscellaneous	28	103½
Total	421	

The miscellaneous group contains the children of parents having occupations which were mentioned less than three times in the testing. Even a group of three is perhaps too small to average as a unit. Doubtless the element of chance had considerable influence in showing the children of policemen to be more intelligent than the children of any other group reported. But when the number of children is large, some significance can be attached to the result. Thus the average I. Q. of 21 children whose fathers are carpenters is 106, while the corresponding I. Q. of 37 children whose fathers work in factories is only 90. It is highly probable that the children of carpenters are, on the whole, more intelligent than the children of factory workers. One would expect the children of ministers and lawyers to be higher in the list, and this would probably be the case if more such children had been tested.

It is interesting to compare the occupational status of the fathers of the children with the results of the American army testing during the war. The army intelligence examination was given to 1,726,966 men. (Army Mental Tests—Yoakum and Yerkes). Each man tested was classified as a member of a group, designated by a letter; the significance of the letters and the percentage belonging to each group of the "white draft" (94,004 men) are as follows:

A	very superior	4.1%
B	superior	8. %
C+	high average	15.2%
C	average	25. %
C–	low average	23.8%
D	inferior	17. %
D–	very inferior	7.1%

Occupational standards based on data for 18,423 men are as follows (Memoirs National Academy of Sciences, Volume XV, Chapter XV):

Occupation	Median Intelligence Rating
Laborer	C–
Miner	C–
Teamster	C–
Barber	C–
Horseshoer	C
Bricklayer	C
Cook	C
Baker	C
Painter	C
Blacksmith	C
Carpenter	C
Butcher	C
Machinist	C
Hand-riveter	C
Telegraph lineman	C
Pipefitter	C
Plumber	C
Toolmaker	C
Gunsmith	C
Mechanic	C
Auto-repairman	C
Auto-engine mechanic	C
Auto-assembler	C
Ship carpenter	C
Telephone operator	C
Concrete construction Foreman	C+
Stock-keeper	C+
Photographer	C+

Occupation	Median Intelligence Rating
Telegrapher	C+
R. R. clerk	C+
Filing clerk	C+
General clerk	C+
Army nurse	C+
Bookkeeper	C+
Dental officer	B
Mechanical draughtsman	B
Accountant	B
Civil engineer	B
Medical officer	B
Engineer officer	A

Thus there is evidently a decided tendency for people to work in positions which require the use of their maximum mental capacity. Since children tend to inherit about the same degree of intelligence as their parents possess, there ought to be differences in the intelligence of children as we proceed from the lower to the higher occupation of their parents. The two following results published by Pintner in "Intelligence Testing" support this view.

In a total of 548 children, the following percentages scored above the group median:

Occupational Group	Percentage
Professional	85
Executive	68
Artisan	41
Laborer	39

The scores of 300 children tested by the Yerkes-Bridges Point Scale are as follows:

Occupational Group	Average Score
Professional	1.42
Travelling salesmen .	1.26
Proprietors, etc.	1.21
Skilled	1.12
Unskilled	.83

It is apparent that children are not created free and equal with respect to their mental abilities. A child's abilities are determined by his ancestors, and environment cannot create new powers or abilities. It then becomes the function of education to measure the inherited capacities of each child and to arrange the environment so as to give full opportunity for the maximum development of all these capacities.

Figure 1 shows the distribution of I. Q.'s of 509 pupils.

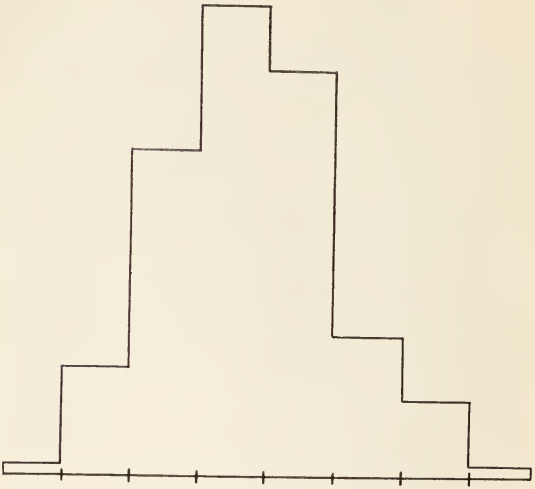
The median, or middle term, I. Q. for the group of 509 would be the I. Q. of the child in the two hundred and fifty-fifth place if they were all arranged in order of magnitude. This median was found to be 96.3.

The average I. Q. of all tested is 97. Theoretically, the average I. Q. of a large normal group should be 100. The occupations of the parents justify us in regarding this as a normal group. Why is the average I. Q. 3% lower than it should be?

American investigators often find the average I. Q. below 100 when using the Terman tests. Recently, many of them have concluded that Terman's tests are a little too difficult for the average unselected child of a given age. Terman regards 16-year-old intelligence by his scale to be that of an average adult. Some psychologists claim that Terman's 14-year-old level more nearly corresponds to the intelligence of the

DIAGRAM 1

Distribution of I. Q.'s of 509 Children



*	60-69	70-79	80-89	90-99	100-109	110-119	120-129	130+
†	4	37	109	154	131	46	22	6 - 509
‡	0.78	7.27	21.41	30.25	25.74	9.04	4.32	1.18 - 100

* I. Q.'s. † Cases. ‡ Per cent.

average adult. The claim is also made that Terman's groups, used in arranging the tests according to age, were not unselected. On page 55 of "The Measurement of Intelligence" Terman gives the distribution of mental ages of 62 normal adults. This group is composed of 30 business men and 32 high school pupils. Most people would regard business men and high school students as being above normal adult intelligence. The results of some investigations bearing on the subject will now be mentioned.

An event which caused much discussion at the time was the testing in August, 1918, of 653 unselected white enlisted American soldiers. Using the Terman tests, the average mental age of this group was 13.4 years.

The following figures give the results of more army testing. (Memoirs National Academy of Sciences, Volume XV.)

	Mean Mental Age	No. of Cases
White draft	13.1	93,965
Colored draft	10.4	18,892
White officers	17.3	15,544

The distribution of the white draft according to mental age is as follows:

Mental Age	Per Cent
16 and over	14.
15	9.3
14	13.6
13	15.8
12	17.
11	12.7
10	7.6
9	4.7

Mental Age	Per cent
8	3.15
7	1.75
6	.2
5	.1
4 and below	.1

These figures show that the mental age of the average white soldier, as judged by the standards of the Terman test, is between thirteen and fourteen years.

Ballard, an English psychologist, using Terman tests in his own country, reports in "Group Tests of Intelligence," that he believes Terman's 16-year-old level as that of average adult to be too high a standard and that the I. Q.'s are accordingly lower than they should be. P. M. Symonds, writing in the *Journal of Educational Psychology*, February, 1923, on the subject, "The Intelligence of the Population in United States," makes this statement:

"On the basis of these figures, . . . it would appear that the median I. Q. of the general population would be about 80, say between 80 and 82, as Terman computes adult I. Q.'s . . . Here there is a reconciliation of the divergent facts that have troubled psychologists these three years. There has been no error. The Stanford revision of the Binet-Simon scale has been standardized on selected individuals. It is true that the average Stanford-Binet mental age of the average man is 13.2 years, or about 13 years."

More evidence on the subject is supplied by Garrison and Tippett in an article, "Comparison of the Binet-Simon and Otis Tests," in the *Journal of Educational Research*, June, 1922. The Otis test has likely been more widely used than any other test. The norms for it are based on a consideration of 25,000 cases, so

should be highly accurate. In the article mentioned, the investigators gave the Terman test and the Otis test to 158 pupils. The median mental age, Terman test, was 12 years 4 months: the corresponding age, Otis test, was 14 years 1 month. Accordingly the authors conclude that the Otis test gives a higher mental age than the Terman test. Thus if the Otis tests are accurate, it would follow that the standard of the Terman tests is too high. Since the norms for the Otis test have been derived from the results of testing more than ten times as many people as were examined to produce the standards of the Terman test, we can reasonably assume that the Otis test standards are more reliable than the Terman standards in regard to the attainments of the average person at various ages.

While carrying on the study, the method was to test all the pupils who were willing in any one class, and then pass on to another class. In all, sixteen classes were tested to reach the total of 509, an average of nearly 32 to a class. The teachers in all cases were teachers of considerable experience. They were each asked to rate the pupils of their classes who were tested as to their intelligence; they assigned each pupil to one of five groups; viz., very superior, superior, average, dull, very dull. Terman and others classify I. Q.'s as follows:

I. Q.	120-140	..... Very superior
	110-120	..... Superior
	90-110	..... Average
	80-90	..... Dull
	70-80	..... Very dull

Comparing the rating of intelligence by the test

and the teachers' rating of intelligence, the results are as follows:

	Test Rating	Teachers' Rating
Very superior	28	11
Superior	46	102
Average	285	275
Dull	109	113
Very dull	41	8
Total	509	509

The two ratings agree well in the average group and in the dull group. The tests placed many more in the very superior and very dull groups than the teachers did. These figures show conclusively that teachers are not good judges of the extreme cases. The bright pupils are really brighter and the dull pupils are really duller than the teachers think them to be. There seems to be also a marked tendency for teachers to regard as bright many pupils who really are only average or dull. This is where the personality and appearance of the child play their parts in influencing the teacher's judgment.

Certain correlations were computed. A correlation is the tendency of two series of measurements to deviate together. Thus height and weight tend to correlate; *i. e.*, the tallest person in a group usually weighs the most and the shortest weighs the least. Of course, there are many exceptions to this, but the tendency is as stated. The degree of this correlation is measured by the coefficient of correlation. Thus in comparing the weights of five people, the rank might be A, B, C, D, E. If the rank, when comparing heights, is also A,

B, C, D, E, the correlation is perfect and the coefficient of correlation is plus 1. But if the rank for heights were E, D, C, B, A, there would be perfect negative correlation, and the coefficient of correlation would then be minus 1. A coefficient of correlation of 0 indicates that there is no apparent relationship existing between the two series of measurements. Thus the more nearly the coefficient of correlation approaches plus 1, the greater the similarity in the ranking of the individual measures in the two series being compared.

A coefficient of correlation of .60 or higher is regarded as highly significant; in practical work, such a correlation would be regarded as evidence that the two series of measures were closely related. When the correlation is below .30, there is little relationship between the two series. A larger negative correlation, −.60 or higher, would indicate a decided tendency for the measures to be in reverse relationship; *i. e.*, if the order of four measures in one series were represented by A, B, C, D, a coefficient of correlation of −.60, or higher, would show a tendency for them to be present in the other series in the order D, C, B, A.

In calculating coefficients of correlation, the method is to arrange the measures in order for the two abilities or traits to be compared. Thus to correlate the ability of a class in spelling with the ability in arithmetic, a list of the pupils in order of merit for spelling would be prepared, and then another list giving the order for ability in arithmetic. There are several formulas used for calculating the coefficient of correlation. In this work the formula used throughout

is that known as the Pearson modification of Spearman's formula. It is

$$r = 1 - \frac{6 \Sigma D^2}{N (N^2 - 1)}$$

where r = coefficient of correlation, D = the difference in rank of a measure in the two lists, N = the number of measures.

The reliability of coefficients of correlation is expressed by the Probable Error (P. E.). The formula for finding this is

$$\text{P. E. } r = .6745 \left(\frac{1 - r^2}{\sqrt{N}} \right)$$

where r is the coefficient of correlation, and N is the number of measures. The P. E. is expressed as a plus and minus quantity. Its meaning will be made clear by an example. If a coefficient of correlation were found to be .58, and the P. E. were calculated to be .095, the result would be expressed as $r = .58 \pm .095$. This would mean that the chances are even that the true coefficient of correlation lies between $.58 + .095 = .675$ and $.58 - .095 = .485$.

In making correlations involving mental capacity, some investigators use the grading according to I. Q.'s as indicating the relative intelligence of the various children in the group. This is usually erroneous because the chronological ages of the children of a group are rarely the same. Thus a boy of 12 years with an I. Q. of 90 has a higher mental age than a boy of 10 years with an I. Q. of 105. In all calculations involving mental ages in this study, the method was first to calculate the mental age of all pupils in a class on a certain day,—usually the day on which the last pupil

of the class was tested. Thus an accurate rating for mental ages was derived.

When a class was tested the teacher was asked to supply a list of the pupils, rating them in the order she thought they should be, based on a consideration of intelligence or native mental capacity only. The examination marks for a long period—usually a year—were also collected and averaged, and a list prepared, giving the order of the pupils based on the results of examinations. Thus for each class there were three lists prepared, giving the order of each pupil in the class based on (1) mental age as determined by the Terman test: (2) the teacher's naïve judgment of the child's intelligence: (3) the examination record for a year's school work.

Correlations (r) and Probable Errors were then worked out from the three lists. The results are given in Table II

Table II shows great variation in the corresponding results for the various classes. The correlation between mental age grading and teacher's grading ranges from .09 to .77. A wide range may be noticed in the other columns also. Thus it is evident that teachers differ widely in ability to estimate the intelligence of their pupils. Also the training and experience of the teacher do not seem greatly to affect this ability. It also seems that the examinations set by various teachers differ greatly in nature. Some are much better indexes of intelligence than others.

The average for column 1 (.41) is not high. It points to the conclusion that teachers' estimates of children's native capacity are significant, but to no marked degree. It also seems certain that factors other than native ability enter into the teacher's judg-

Table II

Correlations of Mental Age Grading, Teachers' Grading, and Examination Grading

Grade	No. of Pupils	Col. 1: Mental Age Grading and Teachers' Grading	P.E.	Col. 2: Mental Age Grading and Examination Grading	P.E.	Col. 3: Teachers' Grading and Examination Grading	P.E.
IX	31						
IX	37	r = .77	± .049	r = .62	± .074	r = .89	± .024
IX	25	.55	± .076	.41	± .092	.82	± .036
IX	8	.68	± .072	.49	± .103	.90	± .025
VIII	35	.60	± .152	.67	± .131	.95	± .023
VIII	25	.09	± .111	.37	± .097	.72	± .054
VII	36	.41	± .112	.44	± .108	.73	± .062
VII	36	.29	± .102	.25	± .105	.62	± .069
VII	31	.43	± .091	.59	± .073	.67	± .062
VI	43	.50	± .090	.37	± .104	.58	± .080
VI	38	.34	± .091	.14	± .100	.21	± .097
VI	43	.25	± .102	.12	± .107	.52	± .079
VI	12	.29	± .094	.18	± .099	.51	± .075
V	24	.40	± .163	.39	± .164	.60	± .124
V	47	.35	± .120	.21	± .131	.34	± .121
IV	38	.19	± .094	.42	± .080	.50	± .073
	509	.47	± .085	.43	± .088	.85	± .030
Average		.41	± .101	.37	± .104	.65	± .064

ment of intelligence. From the average for column 2 one concludes that an intelligence test and a set of school examinations are vastly different things, the correlation being only .37. This raises the question as to whether the school is recognizing and developing intelligence to the degree that it should. A pupil's school success is judged largely by the results of the examinations. Yet the table shows that the intelligent pupil is not the one who leads the class in the examination records. Two explanations are possible. The first is that the school does not make the intelligent boy or girl work up to capacity. The second is that the examinations of some teachers call for the exercise of only a few mental capacities and neglect others. The intelligence test examines the ability in many mental capacities. For success in it a pupil must have a " many-sided " intelligence. Thus it would seem probable that some factors in general intelligence are undeveloped by the schools. The great variety of results recorded in column 2 indicates that the examinations set by some teachers require for success the exercise of much greater intelligence than those set by others. Another fact in support of this tendency is that many boys and girls who are regarded as failures in school become very successful in their work after leaving school. It is highly probable that many of these have mental powers which school examinations never test, but which are essential to success in the world beyond the school. It is also a fact, which will be pointed out and discussed in Chapter VII, that success in work after leaving school depends to some extent on factors which are not tested by either school examinations or an intelligence test.

The average in column 3 is high, as one would ex-

pect. This shows that in ranking pupils for mental ability, a teacher is influenced considerably by the examination record of the pupil. Also in marking an examination paper, the teacher is influenced in her marking by her opinion of the pupil's ability.

CHAPTER II

CLASS-ROOM GRADING

In theory, by our ordinary system of class-room grading, homogeneous groups of children are collected and each group is placed in charge of a teacher. It is recognized that pupils of similar ability work better together, and can be more easily taught by a teacher. It is also recognized that, in general, children are able to profit from instruction roughly in proportion to the degree of mental maturity they have attained. Thus in each class there should be a reasonable homogeneity in the mental ability of pupils who are instructed together.

The classes tested were regarded as being well graded. They are perhaps as well graded as any equal number of classes one might choose at random in any school in Toronto. Let us study the homogeneity of mental capacity shown in these classes. The classes in which less than twenty pupils were tested, were omitted from the study. The greatest mental age and the lowest mental age were ascertained for each of 14 classes: the mental ages were all calculated on the same day so that the results are accurate. Table III gives the results, the ages being expressed in years and months.

Table III

Range of Mental Ages of Pupils of 14 Classes

Grade	No. of Pupils	Greatest Mental Age		Lowest Mental Age		Difference	
		Yrs.	Mos.	Yrs.	Mos.	Yrs.	Mos.
IX	31	17	4	10	8	6	8
IX	37	17	5	10	11	6	6
IX	25	17	6	12	3	5	3
VIII	35	17	5	11	6	5	11
VIII	25	16	10	11	11	4	11
VII	36	16	4	11	4	5	0
VII	36	14	6	9	7	4	11
VII	31	15	9	11	4	4	5
VI	43	14	8	9	8	5	0
VI	38	13	0	9	0	4	0
VI	43	14	1	10	0	4	1
V	24	12	2	9	0	3	2
V	17	12	5	8	8	3	9
IV	38	11	11	7	10	4	1
					Average	4	10.8

Table III shows an average difference in mental age of nearly five years between the brightest and dullest pupils in the classes tested. This is a very great difference to find in a group that is assumed to be homogeneous. How can a teacher effectively instruct, in the same class, pupils whose range of mental age is nearly five years? Can a pupil be expected to work to the best advantage when he is grouped with other children whose mental ages differ from his by three, four, or five years? Yet such conditions seem to be general in most classes. It was shown in the last chapter that teachers do not realize the great range of mental ability in their classes. Intelligence tests have performed a very real service for education by demonstrating the great extent and frequency of individual differences in the mental ability of unselected school children. Schools of the future will have to take account of the great differences in the raw material with which they work.

A further analysis of the mental ages of the pupils was made to emphasize the range of mental ages and also the overlapping in the mental ages of the pupils of the classes. Results are given in Table IV.

Table IV

Analysis of Range of Mental Ages of Pupils in Different Grades

Mental Age	Grades IX	VIII	VII	VI	V	IV	Total
Below 8						1	1
8 to 8-11					1	7	8
9 to 9-11			2	8	10	14	34
10 to 10-11	1		8	33	35	12	89
11 to 11-11	2	1	18	38	20	4	83
12 to 12-11	13	12	37	29	5		96
13 to 13-11	22	17	17	12			68
14 to 14-11	17	14	13	4			48
15 to 15-11	26	11	7				44
16 to 16-11	7	4	1				12
17 plus	5	1					6
Total	93	60	103	124	71	38	489

Some interesting comparisons may be made from a study of Table IV. The 16 lowest pupils in grade IX are evidently not mentally fit for grade VIII; they would have difficulty in doing the work of grade VII. The 8 best pupils in grade VII are mentally superior to the average of either grade VIII or grade IX, while the 10 lowest in grade VII should be put back two grades. Comparing the numbers in all classes, there is a general tendency for those of the highest mental age in any class to be two grades lower than they should be, and those of the lowest mental age in any class to be at least two grades in advance of what their mental age warrants.

In order to obtain more definite results, let us adopt as a standard that a teacher can efficiently teach in a grade all pupils whose mental ages fall within a two-year range. This is certainly not too high a standard. Let us adopt as the standard mental ages for each grade the period of two consecutive years in which the

greatest number of pupils are located. Thus for grade IX the 14- and 15-year groups are larger than the groups of any other two consecutive years. In grade VIII the standard will be the 13- and 14-year groups: in grade VII, 11- and 12-year groups: grade VI, 10- and 11-year groups: grade V, 10- and 11-year groups: and in grade IV, 9- and 10-year groups. All above or below these grade standards are either accelerated or retarded pupils. We then find that a total of 208 pupils, or 40% of the whole number, are either accelerated or retarded. Of these, 88 are accelerated and 120 are retarded. It will be noted that the terms " accelerated " and " retarded," as used here, convey meanings that are exactly opposite to the meanings often attached to those terms in discussing school grading. The accelerated pupils are those who are located in higher grades than their mental ages warrant, while the retarded ones are those whose mental ages are such as to indicate that they could readily do the work of a higher grade. So our net result is that in an investigation of 14 well graded classes, we find 40% of the pupils in wrong classes; 88 of these pupils are in grades beyond their ability and 120 could well do the work of higher grades.

What can be said of the conduct and progress of this 40% group of wrongly graded pupils in the average school?

Let us consider a typical child belonging to the accelerated group. He is with pupils who have more ability than he has, and in time he has the fact duly impressed on him. He does not understand what the instructor is teaching, so he often becomes restless and is regarded as a nuisance. Year after year he finds himself standing near the foot of the class in exami-

nations. Try as he will, he can't get as many marks as other boys in his class; in many cases they are younger than he is. He may try to cheat in a desperate effort to stand on terms of equality with his playmates. If he is caught at such a practice, he is branded as a cheat. If he isn't, there is opened up to him an easy way of overcoming his lack of mental ability: the resultant effect on his character is bad. He probably develops a dislike, often a hatred, for school, and carries with him into the outside world a dislike and hatred of everything having to do with study. The years during which his character is being moulded are spent in unfair competition with those who are mentally his superior, and the evil results of this condition do not develop the type of character necessary for good citizenship.

What can be said of the bright pupil who could do the work of the next higher grade? The work being taught day by day is too easy for him and he becomes bored. He is not encouraged to live up to his possibilities. Since the school work is always easy for him, his will is in danger of becoming flabby from lack of exertion. He becomes conceited, and drifts into the life of a dilettante. How can character develop normally in a child who, during all the years when character is being moulded, never meets a task that calls for his best effort? Many pupils are in this position throughout their school life, and they leave school with false ideas of their own abilities and an undue contempt for the mental calibre of their former schoolmates. In some cases they, in time, become the "peculiar" clever people whom we all meet occasionally.

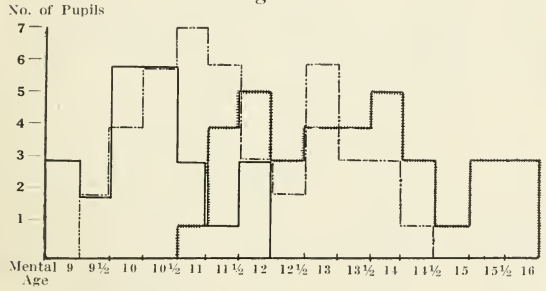
It is the school's duty to place every child in such a grade that he can do the work well, but only by using

considerable effort. Much has been done in America to meet these conditions for the dull pupils. Classes and courses are being constantly prepared for subnormals. But what is being done for the other group? Unfortunately too often the bright pupil is expected to take care of himself. Yet it may be of greater value to society to discover a single gifted child and aid in his proper development than to train a thousand dullards to the limit of their educability. It is through the bright pupils, the coming geniuses, that civilization is advanced. Our future leaders, inventors, and scientists are the bright children of to-day. Yet we are just beginning to pay attention to this group. From an economic standpoint they are the most profitable group of pupils. They will yield a larger return for the money invested to educate them than any other group of pupils. It should be a duty of prime importance for every school to provide the best possible educational opportunities for those who are mentally very superior.

The fact that there is considerable overlapping in the mental ages of one class as compared with the next higher or lower has already been mentioned. This condition seriously handicaps a teacher in her work. To study the matter concretely, reference will be made to the first three classes tested, which happened to be consecutive grades, namely grade VII, grade VI, and grade V. Each grade, theoretically, includes pupils who are a year in advance of those in the class immediately lower. The mental ages of those in grade VII ranged from 16 years 5 months to 11 years 4 months; in grade VI, from 14 years 11 months to 9 years 11 months; in grade V, from 12 years 2 months to 9 years. The lowest pupil in grade VII is then 3 years 7 months mentally

below the best pupil in grade VI. There are 25 pupils in grade VI—over half the class—who are mentally superior to the lowest in grade VII. There are five pupils in grade VI who are above the median age of grade VII (13 years 9 months). Comparing those in grades VI and V, we find there are 19 out of the 24 pupils in grade V who are superior to the lowest in grade VI. There are three pupils in grade V who are above the grade VI median (11 years 9 months). There are even four pupils in grade V with higher mental ages than the lowest in grade VII. Diagram 2 illustrates the overlapping in the three classes.

Diagram 2



——— represents mental ages of Senior Third pupils (grade VI).
——— represents mental ages of Junior Third pupils (grade V).
represents mental ages of Junior Fourth pupils (grade VII).

It is highly probable that overlapping to the extent described in the last paragraph is quite general. As previously mentioned, Bolton Avenue School, in which the testing was done, is regarded by the educational authorities as being well graded. It is an old school in a section of the city where there is comparatively

little changing of addresses. Conditions are uniform from year to year. Accordingly, the grading is probably as good as in most schools and better than is possible in schools where the conditions are not so favorable. Yet from the figures given, one can readily understand that much educational effort on the part of both teachers and pupils is being wasted as the result of inaccurate classification.

Before discussing methods of improving classification in schools, let us consider the comparative ability of the raw material with which a teacher works. Is one class really brighter than another? If so, to what degree? To answer this, the median mental age was found for each class in the cases in which two or more classes of the same grade were tested. All medians were calculated for the same date; namely, January 1. The results are tabulated in Table V.

Table V

Comparison of Median Mental Ages of Different Classes of the Same Grade

Grade	No. of Pupils	Median Mental Age	Range per Grade
IX	31	14-2	
IX	37	14-6	IX -13 mos.
IX	25	15-3	
VIII	35	14-7	
VIII	25	13-7	VIII-12 mos.
VII	36	12-4	
VII	36	13-4	
VII	31	12-8	VII-12 mos.
VI	43	11-7	
VI	38	10-11	
VI	43	12-0	VI -13 mos.
V	24	10-8	
V	47	10-10	V - 2 mos.

Omitting grade V, the average range of mental age when two or three classes of the same grade are compared is one year. If the number of classes of a grade had been larger, the range would undoubtedly have been greater. Thus we see that classes differ a great deal in native ability. The evidence indicates that when two or three of the classes of a senior grade are compared, the difference in mental ability, calculated when the classes are about midway through the year's work, varies to the extent of a year, or approximately 15% of the median mental age of the class.

It has been shown that the method of grading used in most schools produces classes of pupils who differ considerably in native mental ability, and also that there is considerable overlapping of one class when compared with another as to the mental capacity of the pupils. These conditions are met with in most schools in most places. What is being done to improve the condition? Brief accounts will be given of five different methods in operation in various places in an effort to improve class-room grading.

Doubtless the method most frequently used is the multiple-track plan. It is not new and many schools use it in some form. The method as used at present in the schools of Oakland, Calif., will be briefly outlined as a typical example.

A more complete account of the system in practice in the Oakland schools may be found in "Intelligence Tests and School Reorganization," by Terman *et al.*

A study was made in that city of all failures reported in the grades of the elementary schools. All children were then given a mental test and it was shown that approximately 90% of the failures were due to mental inferiority. The schools were then organized

upon a three-track plan adapted to the needs of accelerated, normal, and limited classes, respectively. Not all schools have all types of classes, but each school has the classes, both in type and number, that the mental calibre of the pupils warrants. There are five general types of classes—accelerated, normal, opportunity, limited, atypical. All except the normal classes are permitted to vary the content of the course of study, or the rate of progress of pupils, or both. Accelerated classes are for those pupils who have superior mental capacity. Two plans are being followed with these classes: (1) An enriched course of study with practically normal pace: or (2) increased speed with less enrichment of curriculum. Opportunity classes are for those who have good mental capacity, but because of illness, moving, etc., are working in grades below what they should be; the opportunity class enables such a pupil to "catch up" and then take his place in the class for which he has the mental capacity. Limited classes are for pupils who are so dull or slow that they cannot keep pace with regular class work. The course of study for these classes is modified so that the pupils may pass up through the grades, getting the most essential parts of the work of each grade. Atypical classes are for pupils who are very dull, and who cannot make satisfactory progress in a regular class. Pupils in these classes usually are mentally retarded by three years or more; the enrolment per class is limited to 16 pupils: the course of study differs greatly from that of the other classes, manual work being strongly emphasized. Every attempt is made to organize each school so that each pupil will be placed in the type of class which best suits his needs and his ability. In 1922 the total enrolment of the elementary schools of

Oakland was 26,647. These pupils were grouped in the various classes as follows:

18 atypical classes enrolled 288 ....1. %
57 limited classes enrolled 1,555 .... 5.8%
16 opportunity classes enrolled 303 ...1.1%
Accelerated classes 2,583 .....10. %
Normal classes 21,918 .....82.1%

A corresponding system of grading is carried on in the high schools also.

Regarding the success of the plan, the Director of the Bureau of Research and Guidance of Oakland, states:

"The needs of all classes of pupils can be more fully met at little if any additional expense by a multiple-track system adapted to pupils of superior, normal, or inferior intelligence. This system involves differences in rate of progress through the grades and differences in content of course of study. Experiments carried on in elementary, junior high and senior high schools by classifying pupils according to brightness have demonstrated the feasibility of the plan."

The city of Los Angeles, Calif., is an example of one which seeks improvement in class grading through the agency of "Adjustment Rooms." When a child in an ordinary class-room attracts attention by his failures, he is sent to the school psychologist and examined by him. If found to be below the standard for the grade, the child is sent to an Adjustment Room. He is there given a Placement Test by which the actual level of his development is ascertained. His weaknesses are then explained to him and certain "projects" in the various school subjects are supplied so that he may overcome his weaknesses. He is shown how to select

the work in the various subjects which he needs and then he goes to his seat and works till he has mastered his project. When he feels he is ready for a test, the teacher will give him a project test. If that is passed satisfactorily, he records on his progress card the date on which he passed that particular project. He then begins work on the next one. In the adjustment room there is not ordinary school-room discipline. Each child is there for a definite purpose and all may be working at different levels. A child is allowed to talk, ask or receive help from another child, or consult the teacher, or refer to anything in the room by which he can get help on his project. Each child makes his own daily program, devoting the time to the work he needs most: his program is then submitted to the teacher for approval. The teacher spends about a fourth of the day teaching the pupils as a group such subjects as speed work in arithmetic. During the other three-quarters, each child is really doing supervised study. A very important part of the training is that each child is always competing against himself instead of against some other child of superior or inferior mentality: he tries to improve his record of the previous day. Each success he achieves is immediately rewarded by another step of progress.

Los Angeles now has 52 adjustment rooms and 26 Development Rooms for the feeble-minded. Over 3,000 pupils have been enrolled in the adjustment rooms during the past two years. After a child is sent back to his regular grade, his success from time to time is reported so that it may be known if his weakness or backwardness has been corrected permanently. To date over 90% of those pupils who were in the adjustment rooms have been reported as being

satisfactorily prepared and making good progress.

A third method of improving grading is to grade pupils on the basis of the results of intelligence tests and the teachers' judgment. A city operating this system at present is Miami, Arizona. In that city all pupils from the second to the eighth grade were given group intelligence tests; those in the first grades and a large number of specially selected pupils were given the individual Terman test. These tests showed the great range of mental ages in the classes as previously grouped. The classes were then so organized that the range of mental age was greatly reduced; in most cases it was reduced by more than half. Some few pupils who stood out prominently in the tests were allowed to skip a grade. In carrying out the re-groupings, the results of the intelligence tests formed the basis for the changes made, but in all cases conferences were held with teachers before the final placing of pupils. A new course of study has been provided for these homogeneous classes; the minimum essentials are stated for the poorer classes, and supplementary work and suggested extensions are provided for the brighter classes. Industrial work is prominent in the courses for the very dull classes. The program at present calls for "a careful selection of groups of children with homogeneous mental development as determined by intelligence tests modified by teacher conferences; the diagnosis and standardization of these groups by the use of pedagogical tests; and the application of a flexible and diversified course of study adapted to class groups."

In "Intelligence Tests and School Reorganization" Mr. C. R. Tupper, Superintendent of Schools, Miami, states three lines of improvement of the new system over the old: (1) The percentage of failures

and the percentage of retardation have been greatly reduced; (2) the curriculum and the organization are fitted to the child in place of forcing the child to conform to the system or forcing him out entirely; (3) the program offers opportunity to a large percentage of the children to secure a fundamental and functioning training in industrial and home-making skill which is denied to them under the traditional system.

A fourth method of overcoming the inequalities of grade classification is by grading and promoting the children separately in each subject. Thus, a child may be in the fourth grade in handwriting, the fifth grade in arithmetic, and the sixth in composition. This method is more difficult to work out practically because a child must change grades for the various subjects. It can be done, however, by having all grades in the school teach the same subject at the same time. Then, if all grades of arithmetic are taught at the same time and all grades of spelling at the same time, etc., each pupil can be sent to the grade in which he belongs in each subject.

This method was put into operation at Garden City, New York, during November, 1918. The children were given the Terman intelligence test and also well-known standard tests in such school subjects as arithmetic, reading, composition, and vocabulary. Tables were provided so that the scores in these various educational tests could be translated into year-month indices of maturity. Then a child might have a score which would indicate he had an arithmetic age of 10 years 6 months; a reading age of 11 years 9 months, etc. The next step was to arrange all the children of the school in order from high to low for

each subject tested. Then the pupils were divided into classes according to their accomplishment in each subject. For example, the upper 40 pupils on a list might form grade VIII for that subject; the next 40 would form grade VII, and so on. The child would not be in the same grade for all subjects. The subjects tested were taught in the schools at the same time and the pupils changed grades so as to be in the proper grade for each subject.

In this way homogeneous groups in each subject were obtained. Dr. Raymond Franzen, under whose direction the experiment was conducted, writes as follows, when giving an account of the experiment in "The Accomplishment Ratio":

"Through classification by information and by intelligence we gain a marked increase of attention, concentration, ambition, and other objectives."—*Page 7.*

"Of all the remedial procedure, such as changing teachers and time allotment and books and method, all of which were employed to some extent, it is my opinion that the re-classification was more important than everything else combined."—*Page 31.*

"After two years of such attempts as an ordinary public school will allow, we have removed many of the causes of disparity and increased the association between potential progress and progress in arithmetic, reading and language."—*Page 56.*

Another recent and very promising method of dealing with the problem of grading will be discussed in the next chapter.

CHAPTER III

THE ACCOMPLISHMENT RATIO

Mental Health?

The ability of a child to progress in school work depends on at least four general factors: (1) Mental ability: (2) physical capacity: (3) previous preparation: (4) interest and willingness to work. The purpose of this chapter is to discuss the first and fourth of these factors.

The mental ability of a child is tested by an intelligence test. Such a test claims to measure only the native intelligence. In a school-room it does not follow that the child who is mentally the best pupil is actually doing the best work. Matters of temperament and character; such qualities as carefulness, determination to respond effectively, persistence, energy, all these are important in determining a child's class-room progress. An intelligence test measures such qualities only in so far as they become evident in the factor of general intelligence. The progress of some school children is due to the possession and exercise of these desirable traits to a greater degree than they are exercised by the average child. Thus it is unwise to grade a school on the basis of an intelligence test only.

During the past few years the use of a new measurement has been advocated. This is the Accomplishment Ratio (formerly called the Accomplishment Quotient), which is defined by Franzen as "the degree to which a child's actual progress has attained to his potential progress by the best possible measure of both." The Accomplishment Ratio (A. R.) is found by using standard educational tests in various subjects. By

means of them, the ages attained by a child in various school subjects may be determined. These ages are then divided by the chronological age of the child and the quotients are called his Subject Quotients. The ratio of any Subject Quotient to mental age is called a Subject Ratio, and the average of the Subject Ratios of a child is his Accomplishment Ratio. An example will make the matter clearer. Let us suppose a child of 10 with an I. Q. of 110 is given a standard reading test and makes the score in it usually made by an 11-year-old child. Then his Reading Quotient is 11 divided by 10, or 1.10; this divided by his I. Q. will give his Reading Ratio 1.00. In spelling, perhaps the child makes the score of a 12-year-old. His Spelling Quotient is then 12 divided by 10, or 1.20; his Spelling Ratio is 1.20 divided by 1.10, or 1.09. In language, the child might reach only the 10-year-old level; his Language Quotient is then 10 divided by 10, or 1.00; and his Language Ratio is 1.00 divided by 1.10, or .91. If he were only given the three tests, his A. R. would be the average of the three ratios, 1.00, 1.09, and .91, or 1.00.

Some of the advantages of the A. R. are at once apparent. What a child does is measured in terms of his native ability. The child with a high I. Q. must work just as hard to score an A. R. of 100 as the child with the low I. Q. Each child must work up to his mental capacity to attain an A. R. of 1.00. The aim of each child is to increase his A. R.; he works against his own record and thus ambition and self-reliance are developed. Improvement in A. R. also depends on concentration, interest and effort: such mental factors as these are developed when each child is given the task of improving his A. R.

Accomplishment Ratios were determined for 147 pupils who were grouped in five different classes. Four educational tests were used in doing this; namely, Thorndike Reading Scale Alpha 2, Thorndike Visual Vocabulary Scale, Ayres Spelling Scale, Courtis Arithmetic Test. One reason for choosing those particular tests was that well-known tests of proven merit might be used. A second reason was on account of age norms. Unfortunately, American tests are rarely standardized on the basis of age; usually they are standardized on the basis of grade. To find Accomplishment Ratios, we must have age norms for the tests used. A table by means of which the scores in the two Thorndike tests referred to can be translated into year-month indices of maturity may be found on page 12 of Franzen's "The Accomplishment Ratio." This table was used for those tests. In regard to the other two, the standard scores for the different grades were used, and also a table which appears on page 232, *Teachers' College Record*, May, 1922. The writer of the article assembles in that table six sets of standard age norms for elementary school grades, as given by five authorities; the sets are then averaged with the following result:

Grade	Years	Mos.
III	8	10.7
IV	9	7.5
V	10	11.9
VI	12	5.08
VII	13	7.3
VIII	14	5.6

This table was used, and by means of interpolations

the mental ages corresponding to the various scores in the two tests were calculated.

A word or two of explanation as to the nature of the four tests will be given. The reading scale consists of several paragraphs which the pupil reads; under each paragraph are questions based on it; he writes the answers after reading the paragraph. The point is for him to answer as many questions as possible in the time allowed. The vocabulary test consists of 13 lines of words, 10 words on a line; these words are all names of flowers, animals, games, books, etc.; the child indicates by a letter the class to which each word belongs. Once more there is a time limit to his work. The Courtis Arithmetic Test is made up of examples in addition, subtraction, multiplication and division. A time limit is set, and the score is the number of examples a pupil does correctly in the time. Ayres Spelling Scale contains 1,000 words, graded as to difficulty; a list of 50 of these words, taken from list R of the scale, was given to all the classes tested.

The method was then to administer the four tests to each of the five grades. The answer papers were then marked in accordance with the instructions for each test. The scores were then translated into year-month indices of maturity by the method previously mentioned. The ages thus derived were divided by the respective chronological ages of the pupils on the day on which they were tested; this yielded Subject Quotients (previously called Educational Quotients) for each child in each subject tested. The Subject Quotients were then divided by the respective I. Q.'s, thus producing Subject Ratios. The four Subject Ratios for each child were then averaged, to yield the Accomplishment Ratio. The procedure may

seem rather tedious, but the A. R.'s can be calculated quite readily in practice. One must know only three things to make the calculation; namely, the ages attained by a child in three or four standard tests; his I. Q.; his chronological age.

Tables VI—X give the I. Q.'s, chronological ages, Subject Ratios and Accomplishment Ratios for the five classes.

TABLE VI GRADE VII 26 Pupils

I. Q.	Chron. Age	Read. Ratio	Vocab. Ratio	Spelling Ratio	Arith. Ratio	A. R.
135	12-2	85	95	90	77	87
111	12-2	107	129	96	80	103
111	12-1	103	91	103	82	95
109	13-9	109	111		65	95
107	12-7	84	115	113	90	100
106	12-3		98	117	109	108
104	12-9	70	128	109	99	101
101	12-5	89	111	126	87	103
98	13-7	110	112	118	117	114
98	13-0	115		105	91	104
97	13-3	112	109	115	93	107
97	13-2	110	114	89	102	104
95	13-8	120	113	102	89	106
93	14-2	96	122	115	114	112
92	12-8		121	122	92	112
92	13-5	82	113	115	91	100
90	14-1	108	113	124	77	105
88	14-4	116	132		79	109
88	13-4	122	137	123	80	115
87	15-1	94	84	85	73	84
87	14-0	120	126	117	77	110
87	14-0	112	110	117	96	109

I. Q.	Chron. Age	Read. Ratio	Vocab. Ratio	Spelling Ratio	Arith. Ratio	A. R.
86	15-0	109	115	77	78	95
85	14-1	74		131	142	116
84	14-0	124	113	129	97	116
80	15-1	67	122	123	115	107

TABLE VII GRADE VII 30 Pupils

123	12-2	99	82		69	83
111	13-2	100	133	104	82	104
109	12-1	93	137	115	88	108
108	12-3	83	133		69	95
107	13-5		133	103	68	101
105	13-5		122	96	69	96
99	13-9	100	123	106	71	100
97	13-8	69	92	108	106	94
97	12-8	100	114	106	80	100
96	14-0	74	115	106	83	95
95	13-10	103	102	116	81	100
95	13-6	98	105	103	73	95
95	12-6	113	112	128	101	113
93	13-0	114	111	112	97	108
92	13-9	123	120	122	112	119
92	13-5	113	109	115	86	106
91	14-1	124	116	103	74	104
91	11-10	133		120	119	124
90	15-0	89	101	103	106	100
90	13-6	108	127	108	126	117
89	15-2	91		103	93	96
86	13-0	79	100	122	88	97
85	13-2	100	130	129	100	115
81	15-11	96	125	103	77	100
81	14-4	116	128	117	91	113

I. Q.	Chron. Age	Read. Ratio	Vocab. Ratio	Spelling Ratio	Arith. Ratio	A. R.
79	14-2	88	106	118	85	99
76	14-10	117	106		81	101
74	14-11	89	120	120	100	107
73	13-11	109	99	133	139	120
73	16-7	71	116	82		92

TABLE VIII GRADE VI 39 Pupils

122	11-1	102		101	108	104
118	10-10	125	89	84	92	98
115	11-1	96	98	104	122	105
115	12-9	101	112	104		106
113	11-6	125	111	107	73	104
112	11-8	46	98	106	82	83
112	11-1	83	122	122	107	108
111	11-4	112	123	125	94	113
109	11-10	125	99	114	100	109
109	11-9	85	115	113	74	97
109	12-0	131	122	116	84	113
108	12-1	85	112	107		101
106	11-6	85	125	118	90	104
106	11-6	91	107	119	74	98
106	10-11	52	94	123	131	100
106	10-7	126	134	135	118	128
104	13-3	73	112	106	100	98
104	12-0	118	113	109	99	110
104	11-8	122	97	109	99	107
103	12-10	93	98	87	87	91
101	12-2	92	107	101		100
99	12-7	119	103	104	74	100
99	12-7	118	86	95	96	99
99	12-0		129	128	81	109

I. Q.	Chron. Age	Read. Ratio	Vocab. Ratio	Spelling Ratio	Arith. Ratio	A. R.
99	11-6	97	104		86	96
95	12-11	100	120	113	105	109
94	12-7	120	124	110		118
93	13-8	105	112	69	101	96
93	12-10	86	99	114	108	102
93	12-9	95	130	115	101	110
93	12-2	109	131	120		120
92	14-6	83	105	104	71	91
91	12-5	100	98	112	88	100
87	14-0	91	89	111	76	92
87	12-5	104	124	140		123
86	13-9	112	126	115	85	109
83	13-10	96	109	122	103	107
78	13-2	111	112	121	122	116
71	16-0	113	131	134		126

TABLE IX GRADE VI 13 Pupils

112	11-7	122	122	98	76	104
102	11-3	97	106	111	106	105
102	10-11	100	120	110	85	104
101	11-5	103	113	109	83	102
98	12-1	51	107	112	95	91
94	12-1	108	97	116	84	101
93	11-3	113	137	130	90	117
92	15-1	98	92	94	84	92
88	14-0	80		112	75	89
81	13-10	123	120	131	72	111
79	12-8	112	122	140	113	122
79	13-7	113	116	101		110
77	13-0	111	125	133	96	116

TABLE X GRADE V 38 Pupils

I. Q.	Chron. Age	Read. Ratio	Vocab. Ratio	Spelling Ratio	Arith. Ratio	A. R.
112	11-6	87	104		74	88
109	10-6	105	104		95	101
108	10-8	104	98	115	115	108
108	10-9	100	99	129	87	104
108	10-5	107	121	126	89	111
107	10-6	100	103	94	81	95
105	10-7	99	105	125	87	104
103	10-7	104	103	121	96	106
100	12-6	72	97	97		89
100	11-6	107	104	79	97	97
100	11-7	106	115	120	96	109
100	11-1	125	102	112	86	106
99	13-8	91	98	100	68	89
99	11-8	91	104	112	81	96
99	11-6	105		104	83	97
99	11-0	127	110	105	87	107
99	10-9	111	124	111	87	108
98	11-3	93	115	110	85	101
98	10-9	113	115	139	97	116
97	11-11	111	114	110	82	104
96	12-0	111	85		82	93
94	11-11	110	125	105	87	107
91	12-2	82	126	107	84	100
91	12-7	112	85	71	82	88
91	12-2	108		98	87	98
91	13-1	86	82	102	80	88
90	11-10	115	109	115	97	109
88	12-0	117	123	123	126	122
87	11-10	121	112	92	99	106
85	14-2	100	100	110	80	98

I. Q.	Chron. Age	Read. Ratio	Vocab. Ratio	Spelling Ratio	Arith. Ratio	A. R.
85	13-3	67	95	67		76
83	12-5	124	111	111	90	109
82	11-9	106	119	110	100	109
82	12-11	103	107	100	88	100
81	12-0	112	101	126	123	115
79	15-0	86	105	90	80	90
79	14-6	103	117		90	103
61	15-6	122	139	89	129	120

There is quite a difference in the A. R.'s recorded. The highest is 128 and the lowest 76, meaning that one pupil is accomplishing 128% of what we can reasonably expect him to accomplish, while the other pupil is doing but 76% of what he is mentally capable of doing. As one would expect, most of the A. R.'s are slightly more or slightly less than 100, proving that most pupils use to full capacity the mental ability they have. The ratios for the various subjects are consistent for most pupils. Those in arithmetic are uniformly lower than for the other subjects; doubtless these classes should spend more time at mechanical arithmetic and possibly less in studying the meanings of words. The aim in all should be to have each pupil have a ratio of 100 in each subject. One can find a few examples of pupils who have one very high or very low ratio, indicating special ability, or the lack of ordinary ability, in one subject of study. The subject ratios also indicate the pupils in need of special urging and in what subjects the urging is most needed.

The Accomplishment Ratios are distributed as follows:

A. R.	No. of Pupils
120 - 129	9
110 - 119	26
100 - 109	69
90 - 99	32
80 - 89	10
70 - 79	1
Total	147

If we adopt as a rough standard that an A. R. of from 90 to 109 is satisfactory, there are 101 pupils doing satisfactory work, and 35, or 23.8% of the total, who are doing better than might be expected from their native endowment. There remain 11 pupils, or 7.5% of the total, who are not working up to capacity; they are falling short of what the school can reasonably demand from them.

Schools have been graded on the basis of A. R.'s and excellent results have followed. In doing this there are two factors to keep in mind; namely, the A. R., and the chronological age. If we can group together children of approximately equal age, who have approximately equal A. R.'s, we would have groups which would be truly homogeneous, from the standpoint of ability to learn. In the five classes reported, the proper method to re-classify would be to multiply the A. R. by the chronological age in months for each child. The 30 pupils with the highest products so obtained would form the first class: the next 30 the second class, etc. The five classes would then be fairly homogeneous,—much more so than at present.

In each class the correlation between the I. Q. order and A. R. order was calculated. The results were: —.05,

—.11, —.28, —.34, —.49; the average of the five is —.25. These correlations are fairly high and negative. What is the meaning of the negative correlation? It means that the dull pupils tend to have the higher A. R.'s and the bright pupils tend to have the lower A. R.'s. In other words, the more stupid a child is, the more he tends to get out of education in proportion to his native ability. A glance at any of the tables will show that many of the dull pupils have high Subject Ratios, while the bright pupils have low ones. Hence the pupils who are failing to work up to the limit of their mental capacities are the bright pupils, not the dullards. It is regrettable that the supernormal is neglected and allowed to work below his ability, while the teacher spends much time helping and coaching the dull members of the class. Yet the bright pupils are the more important because they have the power to profit from instruction and to become leaders after leaving school. There should be a shifting of individual attention to pupils on the part of teachers so that every bright pupil in a class is forced to work up to the limits of his mental ability and physical capacity.

The A. R. is probably the best standard by which to judge the efficiency of a teacher to instruct. Bright pupils and dull pupils have exactly the same chance. The teacher who has her class up to 100 or higher in all subjects is certainly doing excellent work. If her teaching of any one subject is weak, the ratios in that subject will be low and thus reveal the weakness. If she has a tendency to teach the class as a unit, without doing much individual work, many of the pupils will have low ratios in one or more subjects. If the A. R.'s of a class at the beginning of a term be known,

at the end of the term they can once more be determined, and the exact progress resulting from the teacher's instruction is apparent.

Many writers advocate using the A. R. to replace the familiar school mark. It would then be necessary to explain its meaning to both scholar and parent. Each child then must work hard to equal or surpass his previous record. If the best pupil in the class has an A. R. of only 90, both the child and his parents know that he is working 10% below the level he should be attaining. On the other hand, the poorest pupil might have an A. R. of 110, thus proving that he is doing even more than can be expected of him, though his accomplishment may seem small when compared with the attainments of the other members of the class. The brilliant child is often over-praised in marks and the dull one not given sufficient credit for his effort. The A. R. as a mark would prevent the tendency to self-conceit on the part of some supernormals and would vastly encourage many struggling dullards, because it would evaluate the accomplishment of the child in terms of his own ability.

If the A. R. became widely used in schools, probably its greatest value would be as an incentive by which teacher, pupil, and parent would become interested in progress. A teacher could prove her efficiency by raising the A. R.'s of her class. The individual child would be very interested in trying to improve his past performance. He would thoroughly understand that he was competing only with himself, and was not trying to beat out pupils of greater ability. If a parent understood the meaning of the A. R., he would be very interested in seeing that it approximated 100, regardless of the standing of the pupil in the class.

If the child's A. R. were 100 and he stood low in his class, the futility of planning an extended high school education for the child would be apparent.

Brief mention will be made of the results obtained by others from the use of A. R.'s.

In order to study A. R.'s, Franzen gave tests in arithmetic, reading, vocabulary, and language to 200 pupils in the Garden City public school in November, 1918, June, 1919, November, 1919, and June, 1920. The average A. R. was 98.24 in November, 1918, and 102.78 in June, 1920. The correlations between I. Q. and A. R. were —.61 in November, 1918, and —.49 in June, 1920. After each examination changes in grading were made so as to raise the A. R.'s when possible. The figures prove that this was accomplished. The correlations prove that the dullards were working harder than the bright pupils. A full account of the experiment may be found in Franzen's "The Accomplishment Ratio." A few conclusions as stated in the book will be quoted:

"After two years of such attempts as an ordinary public school will allow, we have removed many of the causes of disparity and increased the association between potential progress and progress in arithmetic, reading, and language."

"Disparity of school product can be reduced to individual differences in intelligence."

"There is little unremovable difference in industry, conscientiousness and concentration."

"To-day's educational procedure involves a handicap to intelligence."

"The genius has been neglected."

In the *Teachers' College Record*, May, 1922, Murdoch reports an experiment in A. R.'s with 415 pupils

in Honolulu, Hawaii. The pupils were graded on the basis of A. R.'s and very satisfactory results followed. The following quotation is from the article mentioned:

"Certainly the results which we have obtained do show loafing by the brighter pupils and emphasize the necessity for our doing 'forcing' when it is necessary. Our evidence shows that gifted pupils need this most. The backward pupils are much more likely to be accomplishing all that their mental ability warrants."

In the *Journal of Educational Psychology* for October, 1922, Stebbins and Pechstein report a study carried on in Rochester, N. Y. Four hundred children in four grades were given a series of tests and A. R.'s were calculated. The writers conclude their account of the experiment with this paragraph:

"We believe that the Accomplishment Quotient (A. R.) is the fairest and most valuable measure of both the efficiency of the teacher and the pupil, that by reliance on it for guidance, the teacher will come to exact from him that hath even more than he has been giving, and take from him that hath not even less than he has been able to give. And the educational plaudits of 'well done' are seen to be merited more by the retardate possessing the one than by the accelerate endowed with the ten."

CHAPTER IV

COMPARISON OF RESULTS OF TERMAN, OTIS AND TRABUE TESTS

The Otis Group Intelligence Test is an adaptation of the test administered to the soldiers of the American army during the recent World War. Being a group test, much less time is required to examine a class with it than to examine an equal number with the Terman scale. Using the latter, it requires an average of forty-five minutes to test each individual: with the Otis test, a large group of a hundred or more can be examined in the same time. Thus from the standpoint of economy of time, a group test is much preferable: the debatable point is whether the results are as accurate or not.

One hundred pupils who had been given the Terman test by the writer were given the Otis Group Test, Advanced Examination, Form A. All the pupils were members of either grade VIII or grade IX, 48 belonging to the former grade, and 52 to the latter. After the papers had been marked in accord with the instructions for the test, the scores were converted into I. Q.'s by means of a table supplied by Dr. Otis for the purpose. The corresponding I. Q.'s are given in Tables XI and XII. The gains of the Otis I. Q.'s over the Terman I. Q.'s are also indicated; in cases where the Terman I. Q. is the larger, the fact is denoted by a minus sign. The tables are arranged in the descending order of magnitude of the Terman I. Q.'s.

Table XI — Grade VIII

Comparison of Terman and Otis I. Q.'s of 48 Pupils

Terman I.Q.	Otis I.Q.	Gain	Terman I.Q.	Otis I.Q.	Gain
137	159	22	102	124	22
130	149	19	101	127	26
127	130	3	100	101	1
127	122	5	99	108	9
118	110	–8	98	127	29
117	123	6	98	115	17
116	124	8	96	109	13
115	109	–6	96	109	13
112	119	7	94	114	20
109	105	–4	94	97	3
109	118	9	93	94	1
108	132	24	92	106	14
108	127	19	92	100	8
106	118	12	91	119	28
105	142	37	91	104	13
104	113	9	90	109	19
104	116	12	90	100	10
103	109	6	89	116	27
103	118	15	88	111	23
102	125	23	88	101	13
102	126	24	88	108	20
102	99	–3	84	110	26
102	108	6	82	93	11
102	122	20	79	101	22

Table XII Grade IX

Comparison of Terman and Otis I. Q.'s of 52 Pupils

Terman I.Q.	Otis I.Q.	Gain	Terman I.Q.	Otis I.Q.	Gain
120	122	2	98	118	20
119	126	7	98	119	21
118	133	15	97	109	12
117	125	8	96	109	13
116	122	6	95	111	16
114	114		95	87	–8
110	116	6	94	112	18
110	115	5	94	93	–1
109	125	16	92	108	16
108	109	1	92	101	9
108	115	7	92	98	6
107	112	5	92	97	5
106	117	11	90	104	14
106	116	10	90	93	3
104	114	10	89	106	17
104	104		89	104	15
104	103	–1	89	92	3
103	106	3	89	94	5
102	128	26	86	101	15
101	108	7	85	87	2
101	102	1	83	70	–13
101	117	16	83	99	16
100	108	8	82	103	21
99	108	9	78	76	–2
99	110	11	72	73	1
98	109	11	65	70	5

It will be seen that the Otis I. Q.'s are usually higher that the Terman I. Q.'s. In only nine cases are the Terman I. Q.'s higher than the Otis. The average of the hundred Terman I. Q.'s is 99.62, while the average for the hundred Otis I. Q.'s is 110.42. Thus there is an average difference of slightly more than ten per cent. Which scale gives the more accurate results? The Terman indicates that the group examined is practically normal, while the Otis test ranks the group as 10% above normal. There are two arguments in favor of accepting the results of the Otis test rather than the results of the Terman. In the first place, the group should be above average intelligence, as they are all pupils of the highest grades in the public schools. The majority of the really dull pupils never reach these grades. Hence this group should show a higher I. Q. than average unselected children, who, theoretically, average an I. Q. of 100. The second fact is in the derivation of standards for the tests. The Terman standards are based on "the examination of approximately 2,300 subjects, including 1,700 normal children, 200 defective and superior children and more than 400 adults." The norms used in the Otis tests are based on an examination of over 25,000 cases, and should be more accurate than those used in the Terman test.

Garrison and Tippett report in the *Journal of Educational Research*, June, 1922, the results of examining 158 pupils with both Terman and Otis tests. Their median I. Q. for the Terman was 104, and for the Otis 118, thus showing that the Otis test tends to give a higher mental age than the Terman test.

In using the two tests, however, the important thing to know is the degree of correlation between them. Is

the pupil shown to be bright by the Terman test also proven bright by the Otis test? A glance at the tables shows that those with the highest I. Q.'s in one test usually have the highest with the other. The correlations between the order of the pupils based on the Terman test and the order based on the Otis test worked out as follows:

Grade VIII	48 cases	.61	± .071
Grade IX	52	.81	± .032
Average		.71	

Garrison and Tippett report a correlation of .75 for 158 cases. Root, in the *Journal of Educational Psychology*, May, 1922, reports a correlation of .80 for 218 pupils.

These coefficients of correlations are all high and they mean that the relative order of the pupils when examined by one scale corresponds very well with the order when examined by the other. Thus each test tends to confirm the result of the other in indicating the ranking of the pupils in order of intelligence.

It would be desirable to correlate the order in the Otis tests with the order of the pupils based on the teachers' judgment and also on the examination records. Unfortunately many of those who were given the Otis tests were members of different classes. There were only two classes having sufficiently large numbers to give any significance to the correlations when made. The results for these classes are placed, for purposes of comparison, along with the correlations for the same classes between the Terman grading, teachers' grading and examination grading: these latter results are copied from Table II, Chapter I.

Class 1	Grade IX	31 Pupils
Teacher's grading and Otis grading,	r equals .82	± .039
Teacher's grading and Terman grading,	r equals .77	± .049
Examination grading and Otis grading,	r equals .65	± .069
Examination grading and Terman grading,	r equals .62	± .074

Class 2	Grade VIII	35 Pupils
Teacher's grading and Otis grading,	r equals .05	± .112
Teacher's grading and Terman grading,	r equals .09	± .111
Examination grading and Otis grading,	r equals .21	± .107
Examination grading and Terman grading,	r equals .37	± .097

Thus in the first class the Otis correlations are higher than the Terman, while the contrary is the case with the second class. The evidence, though insufficient, would seem to indicate that the Terman test and the Otis test are of approximately equal value in the measurement of intelligence. Garrison and Tippett, in their article referred to previously, conclude as follows:

"The Otis test correlates higher with educational test rankings and teachers' rankings than the Binet-Simon test."

In Chapter II, school grading was discussed. The extent of overlapping of mental ages and the extreme range of mental ages in the various grades, as indicated by the Terman tests, were discussed. The Otis tests revealed similar conditions. In grade VIII the range of mental age was six years one month, the highest mental age being eighteen years seven months, and the lowest twelve years six months. This range is large for 48 pupils who are regarded as having approximately equal mental capacity. The range was even greater for the 52 grade IX pupils. For them it is

eight years seven months; the highest mental age for that grade is nineteen years; the lowest is ten years and five months,—more than two years less than that of the lowest pupil in the next lower class.

A hundred pupils, previously given the Terman test, were also given the Trabue Language Scale B. This is a language test, consisting of ten sentences with blanks in them for which the proper words are to be supplied by the pupils. Instructions regarding marking and scoring accompany the test, and these were followed. Table XIII was then prepared, giving the correlations for the various grades based on the order of the pupils as determined by the Trabue test, the Terman test, the teachers' judgment of the pupils' intelligence, the examination record and the Otis test (two classes only).

One notices the wide variety of the correlations in the various grades: three of them are negative. This is further evidence of the great difference in the ability of teachers to estimate the intelligence of their pupils: also the nature of the examinations set by different teachers varies greatly, and requires the use of intelligence in varying degrees.

It is rather surprising that the correlations between Terman results and Trabue results are so low. It is often claimed that the Terman tests depend too largely on ability in language. If this criticism is valid, one would expect a high correlation between the test and a purely language scale. The evidence here presented does not support the criticism. It is noteworthy, however, that the classes with the larger numbers of pupils have higher correlations between the two tests than the classes with fewer pupils. This leads to the conclusion

Table XIII

Correlations Between Results of Terman Test, Trabue Test, Otis Test, Teachers' Judgments, and Examination Records of 100 Pupils

Grade	No. of Pupils	Trabue and Terman	Trabue and Teacher	Trabue and Exams.
IX	6	.07 ± .273	.74 ± .124	.57 ± .185
VIII	19	.25 ± .145	.18 ± .152	.18 ± .152
VII	21	.35 ± .128	.43 ± .119	.46 ± .116
VI	35	.59 ± .073	.18 ± .109	.09 ± .111
V	19	.10 ± .153	—.01 ± .153	—.04 ± .153
Total	100			
Average		.23 ± .154	.30 ± .131	.25 ± .143

Grade	No. of Pupils	Terman and Teacher	Terman and Exams.	Trabue and Otis
IX	6	.60 ± .176	.67 ± .151	.12 ± .271
VIII	19	.41 ± .128	.44 ± .124	—.06 ± .153
VII	21	.43 ± .119	.59 ± .095	
VI	35	.34 ± .101	.14 ± .110	
V	19	.35 ± .135	.21 ± .147	
Total	100			
Average		.43 ± .131	.41 ± .125	.03 ± .212

that with larger classes the correlations would probably be higher.

The correlations between the Trabue test results and the teachers' estimates of intelligence and examination records are also low:—considerably lower than the corresponding correlations with the Terman test. Thus it is evident that in estimating the intelligence of pupils, a teacher is not influenced to any great extent by the respective language abilities of the various scholars. It also seems evident that language ability does not cause a child to take high marks in the examinations: success in examinations apparently depends largely on factors other than language ability, as measured by the Trabue Scale.

The Trabue test correlates poorly with both the Terman and the Otis tests; these latter two correlate with both the teachers' judgment of intelligence and with the examination records more highly than the Trabue correlates with them. Thus it seems evident from this study that the Trabue test is a poor test of intelligence. It measures only language ability and the correlations given indicate that language ability, as measured by the Trabue Scale, is no great sign of intelligence.

It must be added, however, that this result differs from that of some other investigators, who have found the Trabue test to correlate highly with intelligence.

CHAPTER V

THE CONSTANCY OF THE I. Q.

When a child is given a Terman test and his I. Q. determined, will it remain constant throughout his school life? Some ardent advocates of testing claim that it will. That is, if a child of seven should have an I. Q. of 112, some psychologists state it would not deviate from 112 by more than a point or two, if the child were tested again at eight years, nine years, and so on. Others think that the deviation shown by subsequent tests would probably be much greater. From a practical standpoint, it is very important for the teacher or educator to know the probable limits of possible deviations in the results of successive tests with the same scale.

Thirty pupils were re-tested during the course of the study. It was felt that the number was not large enough to warrant any general conclusions. However, the subject is so important that it seemed imperative to present material from which conclusions might properly be drawn. Hence search was made in various magazines and journals in order to assemble the results of as many Terman re-tests as possible. Various issues of the *Journal of Educational Psychology* and the *Journal of Applied Psychology* supplied nearly all the data found. This material will be presented along with my own, and then conclusions will be drawn.

Six studies dealing with Terman re-tests were carried on between 1915 and September, 1921. In the *Journal of Educational Psychology* for September, 1921, Rugg and Colloton present the results of these studies and draw inferences from the material. The

following table is (in part) copied from the article: the rest of it is deduced from the information in the article.

Table XIV

Results of Re-testing 1,726 Children

Investigator	No. of Children	Interval		Average Difference	Correlation
Terman	436	Less than 1 yr.	86	4.5	.93
		1-3 years	138		
		3-5 years	85		
		More than 5 years	127		
Cuneo and Terman	77	2 days	25		
		5-7 months	21		
		20-24 months	31		
Garrison	62	3 years		4.66	
Poull	126	6 months to 3 years		4.6	
Wallin	61			6.6	
	61			6.2	
	19			6.1	
	120			10.2	
	120			14.1	
Rugg and Colloton	137	10-16 months		4.7	.84
Fermon	233	7 months-4 years		7.5	
Stenquist	274			med. gain 9.5	
				med. loss 5.9	.72
Total	1,726				

The writers of the article state that the average difference is approximately 5 points. For those reported by Terman, Garrison, Poull, Rugg and Colloton —760 in all—the average difference is 4.5. The writers then state as follows:

"This means that the chances are approximately 20 to 1 that the I. Q. of a pupil reported from a single test is within 13 points of his true I. Q. The chances are one in two that an I. Q. from a single test will in-

crease as much as 6 points, or decrease as much as 3; the chances are 1 in 5 that it will increase as much as 12, or decrease as much as 6; the chances are 1 in 20 that it will increase as much as 18, or decrease as much as 9. Hence much confidence can be put on a single I. Q. if the examination is made by experienced and well trained examiners."

Table XV gives corresponding data for other investigations.

Table XV

Writer	No. of Cases	Interval	Average Diff.	Correlation
Burns	77 (borderline)	9 mos. - 4 yrs.	5.3	.91
Gordon	44 (all Hebrew)	1-3 to 2-10	6.8	.84
Lincoln	27		5.25	
Total	148			

Garrison, in the *Journal of Educational Psychology*, May, 1922, reports on 468 re-tests extending over an interval from one to four years. Of these,

8.5% show a difference of more than 10.

89% show a difference of 8 or less.

There is a gain in 55% of the re-tests and a loss in 38%.

The average difference in I. Q. was from 4.5 to 7.2, the grouping being based on age. The higher the age group, the greater the difference usually. The highest age group, 16-year-olds, had the difference of 7.2.

Baldwin and Stecher refer in December, 1922 *(Journal of Educational Psychology)*, to a study involving 143 cases. Their statement is: "The data confirms the findings of the previous study, that for practical purposes the I. Q.'s remain sufficiently constant for a group as a whole, but that the individual

records show fluctuations which are smoothed out in obtaining general averages."

Kuhlmann reports the results of repeated examinations over a period of ten years with 639 feeble-minded children *(Journal Applied Psychology*, 1921). He found the average yearly decline in I. Q. to be

Borderline	2.19 points
Morons	1.21 points
Imbeciles	1.04 points
Idiots	.37 points

In the same volume Root reports two children. One, tested when 9, had an I. Q. of 92. Tested nine months later, the I. Q. was 104. Another child, first tested when 6 years, 8 months old, had an I. Q. of 112; 3 months later the test revealed an I. Q. of 140. The writer explains that these are exceptional cases and attributes the gain to semi-individual methods of teaching which permit direct attack on the peculiar weakness of the pupil; also in both cases, during the interval there was a marked improvement in ability to read and comprehend language.

Teagarten *(Journal Educational Psychology*, September, 1922) also reports two peculiar cases. The children were re-tested after four months. The respective I. Q.'s of the first child were 83 and 90; of the second, 119 and 128. The writer accounts for the increases by the fact that in the first test both children barely failed in certain tests, and in the second test they barely passed in some tests.

Table XVI gives the results of thirty re-tests made by the writer. The interval between the two tests is expressed in months. The differences in the I. Q.'s are expressed as increases of the second test over the

first; a minus sign indicates that the first test yielded the higher I. Q. They are arranged in descending order of magnitude based on the interval elapsing between the successive tests.

Table XVI

Summary of 30 Re-tests

Age 1st Test	Age 2nd Test	Interval	1st I.Q.	2nd I.Q.	Difference
12- 3	15- 5	38	102	108	6
11- 1	14- 3	38	126	115	–11
12- 6	15- 8	38	89	91	2
11- 5	14- 6	37	122	120	–2
12- 7	15- 8	37	97	86	–11
12- 3	15- 3	36	106	103	–3
11- 8	14- 8	36	91	80	–11
10- 7	13- 6	35	98	87	–11
11- 5	14- 4	35	107	98	–9
12-11	15- 0	25	104	108	4
9-10	11- 1	15	101	111	10
14- 2	15- 3	13	86	83	–3
10- 7	11- 8	13	121	111	–10
12- 8	13- 8	12	103	99	–4
13-10	14-10	12	80	83	3
11- 3	12- 3	12	96	104	8
10-10	11-10	12	98	106	8
14- 5	15- 4	11	108	106	–2
13- 5	14- 4	11	77	74	–3
13- 8	14- 7	11	95	87	–8
11-10	12- 9	11	107	97	–10
12- 2	13- 1	11	90	97	7
12- 8	13- 7	11	78	79	1
9- 8	10- 7	11	113	91	–22

Age 1st Test	Age 2nd Test	Interval	1st I.Q.	2nd I.Q.	Difference
12-10	13- 8	10	94	99	5
13- 9	14- 7	10	95	85	−10
12- 9	13- 7	10	86	83	−3
13- 6	14- 4	10	82	76	−6
11- 6	12- 4	10	87	86	−1
15- 3	16- 0	9	99	98	−1

It will be noticed that the tendency is for the I. Q to show a decrease; in 10 cases there is an increase, and in 20 a decrease. The net decrease is 87 points, an average of nearly 3 points per pupil. One pupil, however, accounts for a decrease of 22 points. This child is an example of the odd peculiar case reported by others who have re-tested. In school she is decidedly erratic, and is often a source of mystery to her teachers. At one time she surprises them with the excellence of her work and the cleverness of her answers. At another time she seems almost stupid. Omitting her score, the average decrease per pupil is slightly over two points.

The cause of the decrease is readily understood by looking at the age column for the second test. Eight pupils were over 15 years of age when re-tested; 17 of them were over 14. The Terman Scale needs revision upward, as it stops at the age of 18. A pupil of 15 or 16, when tested, has very little chance of making a high score, because there are few tests for which he can get credit beyond his own age. If a 16-year-old person passed all the tests of the scale, he would be credited with a mental age of 18; his I. Q. would be 18 divided by 16, or 112½. There is no doubt that in the cases reported many of the second I. Q.'s would have been higher if it had been possible to give

more tests beyond the 16- and 18-year-old standards.

Table XVII indicates the differences in I. Q.'s based on the ages at which the second tests were given. It shows that the greatest loss occurred with the 14-year-old group, while the youngest group reported shows a gain.

Table XVII

Age in years	11	12	13	14	15	16
No. of cases	3	3	6	9	7	1
Average dif.	2	−1	−1	−6	−1	−1

The average difference in the second I. Q. as compared with the first is 6.5 points. Omitting the one case which gave a difference of 22, the average difference then becomes 6. The evidence then would indicate that when a group is re-tested, the average second I. Q. is apt to vary from the first by a difference of six points.

The correlation between the respective I. Q. orders for the two testings is .81. This shows that there is but little change in the relative order of the pupils in the two lists. The highest in the first test tend to be the highest in the second test: those who are low in one list tend to occupy similar places in the other list.

Table XVIII indicates the gain or loss in terms of the brightness or dullness of the pupils. Thus the first line of the table means that at the first test there were 3 pupils with I. Q.'s over 120. At the second test, the I. Q. of each of them was lower, the average loss being 7.7 points. The table shows that the bright pupils lost more than the dull ones. The reason has already been given; namely, most of the pupils were quite old at their second test and there were not enough

advanced tests to give them a fair chance of demonstrating their real ability.

Table XVIII

Summary of Losses and Gains in Re-testing of 30 Pupils

I. Q.	Average Gain	Average Loss	No. Gaining	No. Losing
120 plus		7.7		3
110 - 119		22.		1
100 - 109		1.	3	5
90 - 99		2.4	4	6
80 - 89		1.3	2	4
70 - 79		1.	1	1
Total			10	20

The average difference of 6 is higher than that reported in Table XIV for 1,726 pupils. The reason is that of the 1,726 in Table XIV, only 99 were over 12 years of age. Only 4 of the 30 reported by the writer were under 12. As already pointed out, the higher the age above 10 years, the greater the chance of inaccurate calculating of the I. Q. Garrison reports an average difference of 7.2 for 16-year-olds.

Neglecting the feeble-minded children reported by Kuhlmann, the results of over 2,500 re-tests with the Terman scale have been presented. This number is large enough to enable us to note certain tendencies. For unselected children the second test is apt to show a difference up to 9 or 10 points, with an average difference of 5 points. For children of age ten or less, the difference is not apt to be so great as for older pupils. For those above 10, the greater the age, the

greater the possibility of a difference in I. Q. For those above 13, the probability is that the second test will show a lower I. Q. than the first one. There is an occasional child, perhaps one in thirty, for whom the two tests show quite different results. A close study of such a child will usually indicate quite definitely the reason,—certain physical, emotional, or temperamental factors will usually explain it.

In discussing the causes of the differences in I. Q.'s as revealed by re-tests, we must examine the parts played by the three essentials in any test; viz., the one testing, the test itself, the one being tested.

Regarding the one conducting the test, it is understood that he should strictly follow the instructions given regarding the administering and scoring of the test. Any variation in the instructions will likely lead to inaccurate results. In the cases in which the same person conducts both tests, it is not likely that the second I. Q. is influenced to any extent by the examiner.

There are several possible causes of inaccuracies when we come to consider the Terman test,—especially in the higher ages. As already pointed out, it is difficult for a 15- or 16-year-old person to do himself justice on account of the fewness of the tests beyond the 16-year-old level. Up to and including 10 years, all tests have a credit of 2 months; for 12 years, each test has a credit of 3 months; for 14 years, the credit is 4 months; 5 months for 16 years; and 6 months for 18 years. In the first examination a child may just fail in, let us say, two of the 14-year tests and one of the 16. Shortly after, he may just pass these three tests; this would cause an increased credit of 13 months in mental age and likely an increase of about 7 points in I. Q. Many of the tests in the upper years require

for success the passing of two out of three parts; there is no credit for passing one part. It often happens that a child will go along, passing one in several such tests and receive no credit. Another child may pass two in one test, miss all in the next couple, and then succeed in two more of some other test, thus receiving credit. It seems that it would be more accurate to give some credit for each unit in which success is attained. Chance certainly plays a part with the older pupils. The margin between success or failure in a test is often slight; success means an additional credit of 4, 5, or 6 months mental age, while failure means no credit. Another source of inaccuracy occurs in the 16-year-old test. If a pupil passes all the 14-year tests, and all (6) the 16-year-old tests, his mental age is 16 years. If he passes all the 14-year tests and all but one of the 16, his mental age is 14 years plus (5 x 5) months, a total of 16 years 1 month. Thus he is a month in advance by failing to pass one test of those prescribed for year 16. Similarly with the 18-year-old test, a person gains 6 months by failing to pass one test; 6 tests passed yield a mental age of 18 years; 5 passed yield a mental age of 18 years 6 months.

Considering the child, we find many conditions which may cause a difference in I. Q. when re-tested.

If a child is tested and the deficiencies which are revealed are explained to the teacher, it is possible for the teacher, by giving individual attention to the child, to so instruct him that his I. Q. will be greater at a later test. This is especially true if the child be weak in language knowledge, comprehension of what he reads, or simple arithmetic operations. Drill in such

work could conceivably have unintentional but nevertheless indirect effect in improving the I. Q.

A change in the conditions and environment of the child may affect the I. Q. If a child has been weak or puny so that he has been kept at home a great deal, he does not have the range of experiences of the ordinary child. His health may improve so that he enters more fully into the experiences of the average child of his age. His I. Q. will likely show an improvement. A case akin to this is that of the immigrant or one who does not thoroughly understand the language. Such lack of knowledge will cause a low rating; later, when he has acquired a greater knowledge of the language, his I. Q. should be higher.

There is a type of child whose answers show a wide scattering. Thus he may pass all the 8-year-old tests, pass half the 9-year-old, fail entirely in the 10-year-old, and pass some of the 12-year-old. He perhaps has special ability in some type of test which recurs in those for various years. It is always difficult to test such a child accurately. One never knows just when his limit has been reached. There is always the chance of leaving unexplored some part of his general mental ability.

It has now been established that improvement in health and physique will often cause a higher I. Q. Those who work with undernourished children or with those who are physically handicapped have shown by repeated tests that in many cases the I. Q. improves with the health. The girl reported in Table XVI whose I. Q. increased from 101 to 111 in 15 months is an example of such a case. For years she had been weak, sickly, underweight and puny; she was in that condition when first tested. A wonderful improve-

ment in her health took place during the 15-month interval between the tests. There is but slight doubt that her increase of 10 points in I. Q. was due almost entirely to improved physical condition with resultant increased mental capacity.

Some children suffer from nervous disorders which make it very difficult for the examiner to test their intelligence accurately. Such pupils are not at ease and are not doing their best. Many mistakes in answering are made and some of them are due to the nervous condition of the child. There is also the child who is quiet, reserved, bashful, shy, timid, repressed. He will not answer unless he is sure; he refuses to guess. At a later test he may have "come out of his shell" and reveal ability which he did not show the first time. There are also some who give evidence of being rather unstable personalities. Their answers are very erratic; they miss easy tests but have a habit of correctly answering an occasional difficult one. The element of chance seems important with them. Two tests may yield quite different I. Q.'s. Another type is that of the child who is superficial, self-centered, and shows lack of interest and effort. Often he is thoughtless in his answers. An example of such thoughtlessness will be given. In the 14-year-old tests is an arithmetic question in which a man's salary is stated as being $20.00 a week and his expenses $14.00 a week; the question is to tell how long it will take him to save $300.00. Some thoughtless children will do the arithmetical operation correctly, but will answer, "50 days," instead of "50 weeks." A child of this type may be a little more thoughtful and energetic at one time than at the other; this would account for a difference of I. Q.

The writer feels that the emotional condition of a pupil is a great factor in explaining the difference in I. Q.'s in many cases. Especially is this the case with girls at the age of adolescence. Many of those re-tested were girls of 14 or 15 years of age. In many cases they showed a nervous tendency which they could not overcome. As a result they made foolish mistakes. They were extremely anxious to do well; they knew they had a reputation for good work in school and they felt they must live up to it. In one such case the girl shook so much that she nearly fell out of the chair while being tested. It seemed impossible for some of the girls to overcome this condition and, of course, it was impossible for the examiner to make allowance for the mistaken answers which resulted from the condition.

Finally, there is the problem of the possible fluctuation of intelligence. Is intelligence the same at all times? The writer believes that it does fluctuate within certain rather narrow limits. To prove this it is necessary to present facts regarding many successive tests on material of equal difficulty under exactly the same conditions and where the interest and effort of all are at a maximum during every test. The results of certain typewriting tests to be described satisfy these conditions.

As a teacher, among other subjects I teach typewriting to the pupils of a Fifth Form Class. After learning the keyboard and getting certain practice, they are given a weekly speed test. The material used is that supplied by the Underwood Typewriting Company for speed work. Each sheet is of approximately the same difficulty. The tests are given at the same hour each week. After a few tests, the weekly type-

writing test is regarded as the big event of the school week. The average rate per minute is calculated: each pupil knows his record for the previous week and is extremely anxious to improve it. Conditions are the same, the task each week is the same; namely, to write from some previously unseen material as many words as possible in fifteen minutes with as few mistakes as possible,—each mistake causes a deduction of ten words from the total number written. Each pupil is working against his own record only,—not to equal some record which he knows it is impossible for him to attain.

Theoretically, the record should show each pupil to be increasing his speed gradually or remaining stationary for a short time. Appreciable decreases in speed would seem to be due to one of two causes, or to a combination of the two; namely, the emotional attitude of the scholar, or decreased mental ability at certain times.

In Table XIX the results of the work of last year's class for twelve consecutive weeks are given. The test for the first week of all is omitted. The rate per minute for a fifteen-minute period is given in each case. The pupils are unselected: all members of the class who had eleven out of the twelve consecutive tests are recorded.

There are but few pupils who show uniform progress in learning, with no backward slips. Pupils 11, 21, 25 are typical of the ideal learning process:—a little progress, then remaining at a level for a week or two, then another upward spurt. Let us calculate the number who have slipped back 20%, 30%, 40%, etc., in any performance as compared with the one immediately previous. Table XX gives the facts.

Table XIX

Typewriting Record of 34 Pupils in 12 Tests,—Rate in Words per Minute

Pupil	Number of Test											
	1	2	3	4	5	6	7	8	9	10	11	12
1	8	8	11	8	10	11	8	13	15	14	11	
2	6	7	6	6	10	8	15	10	16	17	16	17
3	4	8	10	7	9	10	12	12	14	17	11	15
4	6	6	7	9	3	8	4	9	11	4	2	16
5	4	5	7	5	8	9	11	16	11	13	16	18
6	3	4	8	10	3	7	11	13	13	10	18	11
7	4	4	7	2	3	4	9	10	12	13	5	17
8	3	1	3	6	8	8	8	6	9	8	12	
9	5	2	8	11	7	8	6	14	12	11	16	18
10	5	4	4	11	9	10	9	16	14	18	17	17
11	2	4	4	4	6	8	7	8	9	9	9	11
12	6	9	8	8	11	16	12	16	14	14	18	20
13	3	5	2	2	6	9	7	10	11	9	9	7
14	1	7	8	7	10	12	12	13	14	19	13	11
15	1	3	3	3	4	5	5	5	4	9	6	6
16	0	0	1	0	1	5	3	4	10	6	9	11
17	2	3	4	7	3	10	5	14	11	11	8	10

18	0	4	5	3	5	4	7	6	9	10	13	11
19	0	3	4	3	8	13	10	12	11	16	16	13
20	2	6	5	6	7	9	7	12	13	13	13	
21	7	9	13	13	15	15	18	18	19	19	22	21
22	6	6	4	10	13	10	13	11	14	13	11	
23	5	9	2	11	12	14	18	17	18	15	21	20
24	8	11	13	14	14	21	23	25	18	24	24	27
25	1	4	4	4	5	6	8	8	8	7	11	10
26	1	2	3	6	7	4	10	9	11	13	15	13
27	6	5	4	5	8	9	12	12	11	11	13	
28	7	11	10	14	15	11	9	9	13	6	14	13
29	4	5	6	8	5	8	10	14	11	16	16	14
30	5	6	8	8	7	12	11	9	11	13	15	10
31	4	4	2	4	7	9	12	5	11	8	10	9
32	7	7	6	9	12	13	11	17	17	15	16	18
33	3	4	9	7	9	12	17	12	10	12	5	15
34	5	9	13	13	10	11	11	15	16	15	14	22

Table XX

Summary of Decreases in Rate in Typewriting Record

Number of Times	20%	30%	40%	50%	60%
Once	30	23	14	12	8
Twice	21	10	5	4	1
Three times	10	3	2	1	
Four times	3	1	1	1	

The table means that 30 out of the 34 pupils showed a decrease of 20% in one score as compared with its immediately preceding score at least once: 21 showed a decrease of 20% at least twice: 10 showed a decrease at least three times out of the 12 tests reported: while 3 showed a decrease of 20% four times. There were 23 pupils—over half the class—who showed a decrease of 30% at least once; 14 of the class showed a decrease of 40% at least once.

Thus it is seen that a decrease in score at times is the rule rather than the exception. Moreover, decreases of less than 20% have been ignored. As previously mentioned, conditions for each test were uniform; all were extremely anxious to do their best; all tests were of equal difficulty. The writer believes that the chief causes of the facts revealed are two in number: First, in their anxiety to do well, many become "nervous" and fail to make their best effort: secondly, at times they have greater ability to excel in their work than at other times; *i. e.*, greater intelligence. This latter condition is closely related to the state of the pupil's health, the state of his digestive system, the amount of rest he has had, etc.; such factors as the weather, the ventilation of the room, etc., doubtless affect the results to some extent.

In conclusion a paragraph will be quoted from an article in the *Journal of Educational Psychology*,

November, 1922. The writer is E. A. Lincoln of Harvard University.

"The record which any individual makes on an intelligence examination is not due to his native capacity or natural endowments alone. It is influenced by his mental health, the stage of development of his innate ability, and the general environment, including formal and informal education, to which he has been subjected. Most children are mentally sound and sane, they develop at nearly the average rate, and they grow under the ordinary circumstances of environment. In such cases substantial constancy of Intelligence Quotients may reasonably be expected. This study seems to indicate pretty clearly that when large differences in Intelligence Quotients appear there may be found some reason for them. Thus the discrepancies do not lessen the value of the Intelligent Quotient but increase its usefulness, providing always that a scientific case study of the individual is made."

CHAPTER VI
SIBLINGS AND TWINS

More than four years elapsed from the time the first test was given till all the testing was completed. During that time, in many cases, different members of the same family were examined. Usually the senior members of the family were chronologically older when tested, but in six cases the opposite proved to be true, the junior members being the older. In all, 114 pairs of siblings were tested. There were 39 pairs of sisters; 25 pairs of brothers; 50 pairs included a brother and a sister.

In some cases more than two members of the same family were examined. Table XXI gives the data for six families. The children are arranged in their respective order in their families, the oldest being placed first and the youngest last. The sex and the age when tested are stated.

Table XXI

Summary of the Testing of the Children of Six Families

Family "A"			Family "B"			Family "C"		
Sex	Age	I.Q.	Sex	Age	I.Q.	Sex	Age	I.Q.
Boy	12- 6	127	Boy	12- 5	98	Girl	14- 3	97
Girl	12- 0	92	Boy	11-11	89	Girl	12- 2	88
Girl	11- 8	91	Boy	10- 8	137½	Boy	9- 7	94
Boy	11- 3	99	Girl	10- 1	103			
Boy	9- 3	116						

Family "D"			Family "E"			Family "F"		
Sex	Age	I.Q.	Sex	Age	I.Q.	Sex	Age	I.Q.
Boy	13- 4	127	Girl	11- 8	135	Boy	14- 5	82
Girl	11- 5	104	Girl	10- 9	122	Girl	14- 1	100
Boy	10- 4	108	Girl	9- 2	130	Boy	11- 8	111

Families A and B are recorded because they are the only two cases in which more than three members of a family were tested. The other four families are chosen at random. In Family A, the first child tested is the third in the family, the two eldest children having left school. It is evident that the oldest one tested is the clever child of the family. The youngest is also above the average. Family B is an interesting one. The third child of the family is again the clever one,—in fact his is the highest I. Q. found in all the testing done; the other children are apparently just average. Family C shows three children who are all of approximately equal ability, being slightly below normal. In the case of Family E, the children are all bright and much above the average. Family D shows one bright member and two slightly above normal. The last family recorded shows a wide scattering of intelligence. These examples show that some families have one member who is considerably more intelligent than the others. The general tendency is for the various members to approximate a certain standard of intelligence, as in the cases of C and E. Doubtless there are many exceptions to this tendency, as in Family F.

Table XXII records the differences in I. Q.'s found by testing 114 pairs of siblings. It means that in 30 cases, or 26.2% of all the cases, the difference in the I. Q.'s was between 1 and 5. In 24 cases, the difference was between 6 and 10, etc.

Table XXII

Differences in I. Q.'s of 114 Pairs of Siblings

Differences in I. Q.'s	No. of Cases	Per Cent of Cases
1 — 5	30	26.2
6 — 10	24	21.1
11 — 15	18	15.8
16 — 20	17	14.9
21 — 25	11	9.6
26 — 30	6	5.3
31 — 10	6	5.3
Over 40	2	1.8
	114	100.

This table shows that in nearly 50% of cases, the difference in the I. Q.'s of two members of a family is ten or less. In 78% of cases, the difference is twenty or less. In only 7% of cases does the difference exceed thirty. If we regard a difference in I. Q. of 25 as marking off one member of a family as distinctly brighter than the other members, we see that such a difference occurs in 12.4% of the cases. The general conclusions from the table would be that the evidence shows that when two members of the same family are tested, in one family out of eight the testing shows one child to be distinctly brighter than the other; in one family out of every two, there is very little appreciable difference in the mental abilities of the two children; in one family out of four, there is a noticeable difference in the native ability of one child as compared with that of his brother or sister.

The median difference in the I. Q.'s for the 114 pairs is 10.37. This is approximately twice as great as the median difference between re-tests of the same person, as reported by several investigators in the last chapter.

The older child had a higher I. Q. than the younger in 43 cases; the younger had the higher I. Q. in 71 cases. The average gain in I. Q. of the younger over the older

was 5.1 points. This would seem to show that the younger member of a family is usually more intelligent than the older. Before drawing such a conclusion, however, we must remember a fact pointed out in the last chapter; namely, that an older pupil has not a fair chance of making an accurate score in the Terman tests because there is not a sufficient number of tests giving credit above the age of 16.

Table XXIII gives the ages at which all the siblings were tested; the first line means that 2% of all of them were less than nine years of age.

Table XXIII

Summary of the Ages of 114 Pairs of Siblings when Tested

Age	Per Cent
Below 9	2
9 - 9-11	10
10 - 10-11	18.6
11 - 11-11	21.3
12 - 12-11	17.3
13 - 13-11	16.7
14 - 14-11	6.9
15 - 15-11	6.6
Over 16	.6
	100.

It is thus apparent that 52% of all the siblings were less than 12 when tested; 69% of them were under 13; 86% of them were under 14; only 14% of them were over 14. It seems that all pupils up to and including 13 years of age have an equally good chance of making a high score in the Terman test. If this be the case, we are forced to accept one of two con-

clusions. Either 31% of those tested, who were over 13 years of age, did not have a chance to do themselves justice and thus lowered the score of the older pupils so that it averaged 5.1 points less than the younger pupils' score, or the younger members of the family are brighter than the older. To justify the first of these conclusions, each pupil over 13 years of age must have been so handicapped by the faultiness of the tests that he made an I. Q. which was 16.4 points lower than his real I. Q. One cannot believe that a pupil over 13 years of age is handicapped to this extent. A difference of only 7.2 points in re-tests for a group of 16-year-olds is reported in the last chapter. Perhaps 114 cases is not a large enough number to warrant definite conclusions. If a conclusion is to be drawn, it must be that the evidence indicates that the younger children were more intelligent than their older brothers and sisters.

Table XXIV reports the average difference in I. Q.'s between the younger and older children, based on the actual differences in their ages when they were tested. Thus the first line means that in 24 cases the siblings, when tested, differed in age by less than one year; their average difference in I. Q. was 12.2 points.

Table XXIV

Difference in Age	No. of Cases	Av. Diff. in I.Q.
Less than 1 year	24	12.2
1 yr. - 2 yr.	34	13.8
2 yr. - 3 yr.	32	15.3
3 yr. - 4 yr.	15	10.6
4 yr. - 5 yr.	6	13.6
5 yr. - 6 yr.	1	6
6 yr. - 7 yr.	2	22

Thus the difference is not very much greater when the interval between the tests was long. The results for the last two intervals must be disregarded, as there are only 3 cases included. Thus the gradual increase in the first three groups from 12.2 to 13.8, and then to 15.3 probably represents the result of the inaccuracies arising from the lack of more tests for the higher years. The rate of increase is not large,—approximately 1½% for each year. Accordingly, Table XXIV confirms the conclusion stated previously; namely, that in most cases the younger child actually was more intelligent than his older brother or sister.

As already mentioned, in 50 cases the pairs included a brother and a sister. In 27 of these cases, the girl's I. Q. was higher than the boy's; in the other 23 cases, the boy's I. Q. was the higher. The aggregate gain in I. Q. of the 50 girls over their 50 brothers was 47 points, an average of 0.94 points for each girl. This evidence shows that the tendency is for the girls of the family to be slightly—in fact, very slightly—more intelligent than their brothers.

The average difference in the I. Q.'s of 39 pairs of sisters proved to be 11.5 points: the average difference for 25 pairs of brothers was 13.3 points; for 50 pairs, composed of a brother and a sister, the corresponding difference was also 13.3 points. The results for the pairs of brothers and the pairs of sisters indicate that the range of intelligence in girls is not as wide as in boys. Boys are more extreme; the bright ones are brighter than the clever girls, and the dull boys are duller than the dull girls. The fact that the difference in I. Q. between brother and sister is no greater than between brothers alone is evidence in support of the

theory that one's native intelligence is more the result of inheritance than of training. If training were the more important factor of the two, we would expect to find the resemblance greater between brother and brother, than between brother and sister, but the results show the resemblance to be identical in the two cases.

Two interesting pairs will be noted in detail.

A pair of sisters tested respectively 71 and 119. The younger girl is as bright as her I. Q. would indicate. The older girl seemed very dull. Later she acted in a rather queer manner in her class-room on several occasions. She was examined by doctors, who stated that she was suffering from a form of mental trouble which would ultimately make her insane. Her conduct to date seems likely to verify the doctors' diagnosis. In this family there are the two extreme cases, —one very bright, and the other very dull and facing a period of insanity.

A pair of young criminals were tested. One often reads that criminals are feeble-minded. The pair referred to give every evidence of leading criminal lives and they are far from being feeble-minded. Their I. Q.'s are respectively 92 and 95. The older boy, now about sixteen years of age, has broken into and robbed about 100 houses in the past four years. Arrests mean nothing to him. Attempts to reform him have all failed. He frequently steals and then hides the stolen goods; he does not steal because he needs the articles. His latest escapade was to escape from an Industrial School where he has been confined for some time. He managed to avoid the police for about three weeks. When arrested again, he admitted committing 96 bur-

glaries and pleaded guilty to thirteen charges of housebreaking and robbery; the latter offences were all committed during his three weeks of liberty. Watches, diamond rings, considerable sums of money and other valuables were taken by him. He also found time to hold up a man and rob him of $50.00. He apparently does not belong to a regular gang, as nearly all his thefts and house-breakings are carried on alone. His younger brother is following in his footsteps. Just at present purse-snatching and gambling with the proceeds seem to be occupying his attention. From a long acquaintance and intimate knowledge of both boys, it seems certain that they are confirmed criminals. Crime, just for its own sake, seems to fascinate and delight them. Yet they are practically normal from the standpoint of intelligence.

Five pairs of twins were tested. Table XXV gives the sex, age when tested, I. Q. for each one, and difference in I. Q. for each pair.

Table XXV

Summary of the Testing of Five Pairs of Twins

Pair	Sex	Age when Tested	I.Q.	Diff. in I.Q.
1	Girl	15-8	77	
	Girl	14-7	85	8
2	Girl	14-6	79	
	Girl	14-4	90	11
3	Girl	15-4	81	
	Girl	15-4	80	1
4	Boy	10-8	83	
	Girl	10-8	78	5
5	Girl	9-9	114	
	Boy	9-9	103	11

The average difference in I. Q. for the five pairs is 7.2 points. This difference is but slightly larger than the average difference in re-tests reported in the last chapter. Thus the twins are very similar in intelligence. The third pair are practically equal, but one point separating them. The physical resemblance in the case of the three pairs of girls was extremely marked. It was almost impossible to tell one from the other. In their answers to the respective tests, the similarity of response was also very noticeable. For example, the first pair both showed a tendency to answer slowly and to miss most tests which required an answer within a specified time; they were also poor with questions requiring the use of reasoning ability.

The I. Q.'s of all except the last pair are low. This suggested two possibilities; either the twins came of families of low mental calibre, or twins tend to have lower I. Q.'s than children who are not twins. To study these possibilities, other members of the same families were tested when possible. The first pair are the only children of the family: the second pair have only one older brother whom it was impossible to test. The third pair have a brother who, aged 8 years 8 months, showed an I. Q. of 100 when tested. The fourth pair have an older sister whose I. Q., when 13 years 7 months old, was 79. The last pair have two sisters, respectively, age 10 years 4 months and 6 years 7 months when tested: their respective I. Q.'s were 133 and 127. Thus the average I. Q. of the three pairs of twins having brothers or sisters who were tested was 90: their four brothers or sisters had an average I. Q. of 110. This difference of 20 points is significant. One cannot generalize when dealing with so few examples.

Nevertheless, it seems highly probable that twins are less intelligent than other children of the same parents. For example, the fifth pair reported are both above the average in intelligence, but their sisters' I. Q.'s are 133 and 127, causing them to be ranked as very bright. It is quite probable that twins often do not get as good nourishment, either before birth or in infancy, as a single child would get. If there is any handicap in the environment of children who are twins as compared with that of children who are not twins, the twins are the ones who suffer from such a handicap. This may account for the mental difference found. A close study of two or three hundred pairs of twins would be very interesting, and would doubtless reveal some facts at present not generally known.

CHAPTER VII

THE RELATION OF INTELLIGENCE TO EARNING POWER

From a financial standpoint, is intelligence worth while? Is the salary earned by any individual closely related to the intelligence possessed by him? Does the person with the higher I. Q. earn more money than the one with the lower I. Q.?

An attempt was made to answer these questions. As tested pupils left school and started to work, they were asked to report from time to time the nature of the work they were doing, the salary earned, etc. Recently all such pupils were asked for a final, up-to-date report. The information gleaned from these statements forms the material for this chapter.

There were many difficulties encountered in securing accurate information for this purpose. Obviously, such a study must extend over a period of years. It often takes a boy or a girl a year or two to prove his worth to any firm. But pupils are lost track of as time goes on, in spite of every precaution to guard against such a loss. Some move to unknown addresses; some leave the city; others, possibly through illness, stop work and remain at home for a time.

Two classes of pupils were studied. The first group is composed of those who all had practically the same training and left school to go to work of the same kind. As a teacher, I am employed in teaching a Commercial Class. The pupils, in general, are those who have just completed the Public School Course. After studying commercial subjects for a year, or in some

cases two years, they take office positions. This group is an excellent one to study because all have had the same training and start out in business life with practically equal opportunity.

Table XXVI gives the information for 37 pupils of this class. The annual salary, based on 52 weeks or 12 months, is given in each case.

Table XXVI

Salaries of 37 Graduates of a Commercial Class

Pupil No.	I.Q.	First Salary	Months Employed	Present Salary
1	94	$520.00	42	$ 832.00
2	96	364.00	24	1,092.00
3	92	780.00	24	988.00
4	94	520.00	24	728.00
5	89	520.00	24	676.00
6	107	624.00	22	780.00
7	101	416.00	21	572.00
8	114	416.00	20	728.00
9	116	416.00	20	780.00
10	99	468.00	19	832.00
11	89	468.00	19	624.00
12	89	390.00	18	686.40
13	92	468.00	16	624.00
14	108	520.00	15	624.00
15	94	468.00	14	572.00
16	106	780.00	12	780.00
17	98	520.00	12	676.00
18	110	520.00	12	624.00
19	98	364.00	10	572.00
20	103	364.00	10	364.00
21	117	660.00	9	720.00
22	119	416.00	6	468.00

Pupil No.	I.Q.	First Salary	Months Employed	Present Salary
23	78	364.00	4	468.00
24	109	468.00	4	468.00
25	90	520.00	4	520.00
26	92	660.00	3	660.00
27	99	624.00	3	624.00
28	91	624.00	2	676.00
29	104	780.00	2	780.00
30	99	520.00	2	520.00
31	108	780.00	1	780.00
32	119	780.00	1	780.00
33	89	720.00	1	720.00
34	91	416.00	1	416.00
35	92	500.00	1	500.00
36	108	468.00	1	468.00
37	98	450.00	1	450.00

In correlating I. Q. grading with present salary grading, it is evident that some allowance must be made for the length of time the pupils worked. Accordingly they were divided into three groups: First, those working eighteen months or more; secondly, those working nine months to sixteen months, both inclusive; thirdly, those working less than nine months. Correlations were then worked out between the I. Q. grading and the grading on the basis of the first salary earned; also between the I. Q. grading and the grading based on the present salary earned. Table XXVII gives the results.

The average correlations are not high, but they indicate that, when pupils first secure positions, the relation between native ability and salary earned is not high. Gradually, ability is recognized and paid for,

Table XXVII

Correlations Based on Ranking According to Salaries Earned

Group I — 12 Cases	Group II — 9 Cases	Group III — 16 Cases	Average
I. Q. and present salary, r equals —.05 ± .194	I. Q. and first salary, r equals .61 ± .136	I. Q. and first salary, r equals .17 ± .163	r = .24
I. Q. and present salary, r equals .27 ± .180	I. Q. and present salary r equals .44 ± .180	I. Q. and present salary, r equals .20 ± .161	r = .30

r between I. Q. grading and present salary for 37 equals .24 ± .104

though the correlation is not as high as one might expect. One reason is doubtless the time element. In group I, r changes from −.05 to .27 in eighteen months. If this rapid rate of re-adjustment continued for a five-year period, the relation between ability and salary would be very high.

There are several facts which help to explain the comparatively low coefficients of correlation. There is, first, the fact that a pupil of average ability may become somewhat of an expert in one branch of knowledge, and he is paid accordingly. For example, the boy who receives the next to highest wage, having an I. Q. of only 92 (No. 3 in table) is an expert penman. He works entirely at bookkeeping and his special ability is recognized. The element of personality doubtless is an important one in advancement. Several of those whose salaries are high have much more attractive personalities, but no greater intelligence, than those with lower salaries. The possession or lack of desirable qualities and habits is evidently an important consideration. A third fact influencing salary is that, while all pupils received a similar business training and prepared for office work, 10 out of 37 do not work at office work. Two girls became telephone operators, one a milliner, one a factory worker, one boy learned the printing trade, another the butcher business. This is a great waste of training—27% of those trained made no use of the training received. One wonders if this condition is as widespread as in the cases studied. If so, the need for vocational guidance, based on a study of the child's ability and aptitudes while he is still in the public school, is emphasized. It is certain also that the element of chance has played a part in determining the

salaries of those studied. The importance of this factor would tend to disappear in a longer time.

It will be noticed that no mention has been made of the sexes. Are boys paid more than girls? The group is made up of 11 boys and 26 girls. The average of the first salaries earned by the boys is $500.00; the girls' average is $544.46. This leads to the surprising conclusion that boys trained for office work at first earn less money than girls similarly trained. The averages of the present salaries are, for the boys, $687.67, and for the girls, $638.00. Thus it is apparent that the boys' salaries increase more rapidly than those of the girls. The average of all boys' salaries is $593.83; of all girls' salaries $591.23. These averages are so nearly equal that there seemed to be no reason for separating the sexes in calculating the correlations.

The average I. Q. of the group is 99.78. Thus it seems that the average boy or girl training for office work is a person of average ability. He can apparently succeed with an I. Q. of 100. It is noteworthy that the average of the I. Q.'s of the ten who do not do office work, or work related to their training, is 94. In some cases possibly lack of sufficient ability caused them to try different work. The average first salary of these ten is $468.80; the average present salary is $560.80; the average of these two sums is $514.80. Comparing this last sum with the corresponding average of their former classmates, it is apparent that these ten people are earning an average salary which is over $75 less than the average salary of those who are employed at the work they were trained to do.

The second class of pupils studied is composed of

those who just dropped out of school and, in most cases, "got a job." Some few of them received special training in some branch of knowledge. There are 36 in the group, 25 being boys. Of the 36, 14 had passed the Entrance examination, but only 3 had attended a secondary school; several attended night schools and took various courses. The following table gives information about them, including the nature of the work they are now doing.

Table XXVIII

Summary of Salaries and Occupations of 36 People

Pupil No.	I. Q.	First Salary	Months Employed	Present Salary	Occupation
1	73	$624.00	54	$936.00	Sales clerk
2	80	364.00	48	702.00	Telephone operator
3	106	520.00	42	832.00	Salesman
4	82	468.00	40	720.00	Clerk
5	97	728.00	38	884.00	Telephone operator
6	115	572.00	36	676.00	Clerk
7	88	416.00	36	468.00	Cashier
8	84	450.00	36	600.00	Truck driver
9	92	520.00	36	920.00	Garage man
10	98	416.00	35	468.00	Factory work
11	87	468.00	33	650.00	Printing
12	91	416.00	30	572.00	Electrician
13	82	420.00	28	720.00	Drafting
14	93	312.00	24	468.00	Clerk
15	96	360.00	23	572.00	Advertising
16	102	520.00	22	780.00	Clerk
17	127	312.00	21	468.00	Messenger
18	128	364.00	20	624.00	Men's furnishings
19	100	480.00	19	676.00	Stenographer
20	104	624.00	19	676.00	Stenographer

Pupil No.	I.Q.	First Salary	Months Employed	Present Salary	Occupation
21	88	600.00	18	960.00	Operator
22	85	416.00	14	520.00	Stock-keeper
23	99	312.00	13	416.00	Clerk
24	83	624.00	11	728.00	Telephone operator
25	76	468.00	11	780.00	Factory work
26	80	468.00	10	728.00	Factory work
27	85	416.00	9	416.00	Clerk
28	98	468.00	9	780.00	Butcher
29	99	468.00	6	520.00	Assembling
30	70	416.00	6	416.00	Messenger
31	80	364.00	6	364.00	Messenger
32	83	424.00	4	424.00	Factory work
33	81	416.00	4	416.00	Factory work
34	96	775.00	2	775.00	Salesman
35	89	520.00	2	520.00	Factory work
36	89	624.00	1	624.00	Factory work

It will be noticed that the pupils in this group have worked, on the average, a longer time than the pupils of the first group. For purposes of correlating with I. Q. grading, they were divided into three groups: First, all working 30 months or more; secondly, all working 11 months to 28 months, both inclusive; thirdly, all working less than 11 months. Correlations were then worked out as in the first group; results are given in Table XXIX.

The coefficients of correlation here do not show any constant tendency. This is not surprising, as the salaries are complicated by the fact that some are employed at work of such a nature that they reach a high salary, nearly the maximum for that work, in a short time. Others are learning some business; the salary for the first few years is low, but will later increase.

Table XXIX

Correlations Relating to Those Recorded in Table XXVIII

Group I — 12 Cases	Group II — 13 Cases	Group III — 11 Cases
I. Q. grading and first salary grading, r equals .34 ± .171	I. Q. grading and first salary grading, r equals —.12 ± .184	I. Q. grading and first salary grading, r equals .71 ± .100
I. Q. grading and present salary grading, r equals —.29 ± .178	I. Q. grading and present salary grading, r equals —.34 ± .183	I. Q. grading and present salary grading, r equals .67 ± .101
Months worked and present salary, r equals .70 ± .099	Months worked and present salary, r equals —.13 ± .183	Months worked and present salary, r equals —.01 ± .203

r between I. Q. grading and present salary for 36 equals .11 ± .110

Many of the apparent inconsistencies can readily be accounted for by one who knows each person intimately and the work being done. The highest salary quoted, $960.00, is earned by a girl who operates a stock quotation machine. The work calls for great accuracy at all times. To quote from her account of her work: "The slightest mistake means a blue tag. Two of these in a month means that you have to see the manager. If it proves to be your fault a suspension of a day or two follows." The next highest salary reported, $936.00, is earned by the boy who has worked the greatest number of months; length of service accounts for this salary. The third largest, $920.00, is explained by the nature of the work being done; he can never expect to earn a great deal more at his work. The boy with the highest I. Q. is learning the men's furnishing business. He is with a wholesale house and his financial reward will undoubtedly increase with time. The boy with an I. Q. of 127 is bright, but very careless and indifferent. As yet, neither he nor his family have shown any ambition for him to succeed or advance in either school work or business.

The discussion of individual cases might be continued. Certain general conclusions are apparent. Usually the boy or girl who was well thought of in school earns a high salary: the converse is equally true. The importance of a pleasing personality and the possession of attractive and esteemed qualities is very evident. In the case of those in group I who have worked the greatest length of time, the tendency is for the salary to depend very largely on the time the pupil has worked (r equals .71). This is not the case with the other two groups.

For purposes of comparison the average salaries were ascertained, as follows:

Average boys' first salary	$457.00
Average boys' present salary	616.68
Average girls' first salary	518.91
Average girls' present salary 671.10	
Average first salary of all who had passed Entrance examination 507.93	
Average present salary of all who had passed Entrance examination 723.36	
Average first salary of all who had not passed Entrance 449.10	
Average present salary of all who had not passed Entrance 581.16	

These figures show the following tendencies.

1. Girls for the first four years after leaving school earn more than boys.

2. The average first salary of those who had passed the Entrance examination exceeds by $58.83 the corresponding salary of those who had not done so. The difference between the corresponding present salaries is $142.20. It thus appears that those who have completed the Public School Course start at a higher salary than those who have not, and also the salaries of public school graduates increase more rapidly than those of pupils who have not completed the Public School Course.

CHAPTER VIII

THE FACE AS AN INDEX OF INTELLIGENCE

Some people are quite sure that they can make a fairly accurate estimate of a person's intelligence by looking at his face. One often sees articles in magazines and newspapers giving information to guide one in making such judgments. The following quotations are taken from a recent newspaper article:

"If the face is broad across the forehead, the reasoning ability is excellent, and the further back the ear is from the face the more brain capacity is shown."

"A certain 'squareness' about the lower part of the chin indicates strength of mind."

"The happy medium is a fairly broad or oval face, with a wide forehead and slightly showing cheekbones, a square chin, or inclined to be square, and the ear set far back. This nature will have a well-balanced mind, sound common sense, and much reasoning ability."

The relation between the appearance of the face and the degree of intelligence has been scientifically studied by means of photographs. The method has been for an investigator to procure photographs of people whose intelligence has been determined by means of intelligence tests, and then to have competent judges estimate the intelligence of the people by studying the photographs. The result of such studies has been to show that people are not able to estimate intelligence by a study of photographs to any marked degree.

A typical study of the kind is that described by

L. D. Anderson in an article entitled "Estimating Intelligence by Means of Printed Photographs," which may be found in the *Journal of Applied Psychology*, June, 1921. This study will be described in detail. Investigations in this field have also been carried on by Pintner and Hollingworth.

Mr. Anderson states that his objects were (1) to determine the reliability of photographs for indicating the intelligence of strangers; (2) to seek out differences in the respective ability of various persons to judge intelligence by this means. In his study he used the photographs of 69 employees of a commercial company. These employees had all been given an adaptation of the army intelligence test and the scores in this test were used as the criterion of intelligence. The 69 photographs were rated for intelligence by twelve judges, selected from a group of graduate students and instructors in psychology. They were asked from a study of the photographs to pick out the 7 most intelligent, the 7 least intelligent, 14 superior people, not as good as the best 7, and 14 inferior, not as poor as the lowest 7. This left an undifferentiated middle group of 40%, only the upper 30% and the lower 30% being required.

The correlation between the assigned ratings and intelligence as measured by the test was plus .27. Mr. Anderson concludes: "It seems doubtful whether for practical purposes intelligence can be gauged by a study of a man's photograph." He also noted that several of the judges picked out a particular feature, such as the lustre of the eye or the width of the brow, from which to draw their conclusion.

The present writer undertook a study by means of photographs in an effort to answer two questions: (1)

Is there a type of face which indicates the possession of great intelligence, or the lack of ordinary intelligence? (2) If so, is the average observant person able to identify such types from photographs?

The method of studying the subject was by means of composite portraiture, as described in Galton's "Inquiries into Human Faculty." A composite portrait was made of eleven of the brightest pupils tested; their I. Q.'s were respectively, 137½, 135, 130, 130, 129, 127, 126, 123, 123, 122, 121; there were seven girls and four boys. Another composite portrait was made of eight of the dullest who were tested; there were four boys and four girls; their respective I. Q.'s were 75, 74, 73, 73, 72, 72, 69, 61. The two portraits were then compared.

The steps in making the composite portraits were as follows: (1) All the members of a group for one picture were taken to the photographer together. Each one was photographed under the same conditions of light and shade, in the same aspect, and so that the eyes would be in the same position on each plate. (2) The plates were developed and prints made in the ordinary manner, all prints being of the same size. (3) The prints were then superimposed like the successive leaves of a book, so that the eyes of each picture lay exactly in front of those of the one behind it; the other parts, such as the mouth or nose, would also be very nearly in front of the corresponding parts of the next lower picture. (4) The prints of the group were then fastened against an upright board, so that each one could be exposed flat and then removed. (5) The camera was then set up to take a picture of the prints. (6) The prints were photographed, one at a time, each

one being exposed for exactly one second. Thus an image of each print was thrown in succession on the same part of the plate in the camera, for the same length of time.

The effect of composite portraiture is to bring into evidence all the traits in which there is agreement, and to leave but slight trace of individual peculiarities. Each of the prints used in making a composite portrait contributes something to the final result, which is not the picture of any individual, but rather the portrait of a type. Traits peculiar to the individual will disappear or leave a blur in the final product.

By means of such composite portraits there should be produced a type of intelligent face and a type of dull face. The portrait of the bright children was obtained from the faces of eleven of the brightest in an unselected group of 509 children; the other portrait was obtained from the faces of eight of the dullest in the same group. Thus if intelligent people have some common facial traits, or if dull people have common facial characteristics, the composite portraits should emphasize the characteristics that are common in each of the respective groups. Portrait 1 is that of the eleven intelligent children; portrait 2 is that of the eight dull children.

Before making comparisons, we must consider the photographical accuracy of the portraits. The one photographer did the work for both under the same conditions; the pictures were even taken at the same time of day. The photographer states that one picture is photographically as accurate as the other.

The average of the I. Q.'s of the children contributing to the first portrait is 128, and the average for

Portrait 1

Portrait 2

the second is 71. The difference in the faces does not seem to be proportional to the great difference in intelligence. One notices at once that the face of the bright pupils is longer in proportion to its width than the face of the others; the eyes are larger and nearer each other. The lower part of the face of Portrait 1 is not as pleasant as that of Portrait 2. There are two facts to account for this. In the first place, the distance from the eyes to the mouth in two or three of the children in the first portrait was slightly greater than the corresponding distance in the other faces. Since the eyes of all were superimposed, this slight difference would cause these two or three mouths to be slightly lower than the others. Secondly, two sisters, who are in portrait 1, have mouths of a very peculiar appearance and they probably influenced the appearance of the mouth of portrait 1 considerably.

The answer to the first inquiry of this investigation is, then, that the degree of intelligence is not shown to any marked degree by the types of face. When comparing the faces, people may differ vastly in their judgments, as will be shown a little later. The two most obvious and indisputable differences are that the type of intelligent face is "long-featured", while the type of dull face tends to be more "round-faced"; also, the eyes of the bright type are larger than those of the dull type.

In order to test the ability of people to judge the type of intelligence from a study of the photographs, the portraits were shown to more than 600 people. Some 20 of them refused to give an answer, so judgments were recorded for 590 people. Of these, 420 were teachers in the high or public schools of Toronto:

132 were business people, many of them being employers of labor: 38 were medical doctors. The three groups—teachers, business people, and doctors—were chosen because they include those who are making judgments of the intelligence of others from time to time; yet the purpose in making such judgments differs with the respective groups. Collectively, they should represent the ability of competent, observant judges. On being shown the photographs, each person was told that they represented different types of intelligence, one being more intelligent than the other; they were asked to study the portraits and then to indicate which they judged to be the more intelligent, and also to state whether they considered the difference in intelligence to be great or small. Any comments made by the judges while studying the photographs were noted. The results are given in Table XXX.

The summary means that 20.5% made the correct judgment; namely, that Portrait 1 indicates much greater intelligence than Portrait 2. 28.3% judged No. 1 to be slightly more intelligent than No. 2. Practically as many made an entirely wrong judgment as an entirely right judgment, the respective percentages being 20.4 and 20.5. The judgment most frequently made was that No. 2 was slightly more intelligent than No. 1. The two totals, 288 for No. 1 and 302 for No. 2, show that the result is about what one might expect from a chance selection,—nearly 50% for each. Thus it is evident that if there is a type of face which indicates the possession of great intelligence or the lack of ordinary intelligence, competent, observant judges are not able to recognize it.

It is interesting to note the comment made by the

Table XXX

Summary of 590 Judgments of the Intelligence Revealed by Two Photographs

Occupation	Portrait 1			Portrait 2		
	Judged No. 1 much more intelligent	Judged No. 1 slightly more intelligent	Total	Judged No. 2 much more intelligent	Judged No. 2 slightly more intelligent	Total
Teachers	66	126	192	94	134	228
Business People	47	28	75	24	33	57
Doctors	8	13	21	2	15	17
	121	167	288	120	182	302
Per cent	20.5%	28.3%	48.8%	20.4%	30.8%	51.2%

judges. The following comments were made by medical doctors: The mouth of No. 1 is cruel; No. 1 has a temper and No. 2 is very placid; the eyes of No. 1 are brighter; No. 2 has a more kindly nature. One doctor remarked that he did not like the slanting eyes of No. 2. The first comment made by another doctor, a throat specialist, was that he would like to remove the adenoids from No. 2. A nurse remarked that she thought No. 1 had adenoids.

The remarks made by many teachers indicated that they were considering the character and disposition more than the intelligence. Some said that neither face was intelligent; others maintained that both were below normal intelligence; a few thought that the pictures were of one and the same person and that there was no difference in intelligence. One teacher claimed that the pictures show the dispositions rather than intelligence.

Comments made by teachers about No. 1 were as follows: The forehead is better; the distance from ear to ear indicates greater intelligence than in No. 2; the mouth is firm but sulky; the lines of the lips are hard; the face is sour looking; the face is bad, being that of a criminal; the upper part of the face is intelligent, but this is contradicted by the mouth; steady, keen, sharp eyes; firm jaw; finer features; the face is that of a schemer; the face shows more stability than No. 2; there is a lack of poise; a stony, vacant stare; the nose and the set of the eyes show the face to be that of a mental defective.

Typical comments regarding No. 2 were as follows: The mouth is more pleasing and less sullen; the smile gives the face a silly look; the chin is better; the face is more pleasant and cheerful; forehead is better; the

general shape of the head is better; the eyes are better; a better defined nose; a more strongly defined jaw and nose; the eye of No. 2 is more that of a student,—also more indifferent and lazy. Several noted the difference in the size of the eyes.

The comments made by the business people are included in those made by the teachers. More of the business people, however, seemed to judge by the eyes, and this probably accounts for the fact that the business people were better judges than either the doctors or the teachers.

The confusion in the comments noted indicates that people interpret the same feature in quite different ways; to some, the eyes of No. 1 stamp the face as being far more intelligent than No. 2, while others regard the eyes of No. 2 as indicating superior intelligence over No. 1. Hence, any face might be judged to be intelligent by some people and unintelligent by an equal number of other competent judges. Thus it is evident that personal judgments of the intelligence of others, when based on the features of the face, are very unreliable.

CHAPTER IX

CONCLUDING NOTES

Any teacher knows that the educable capacity of some children is higher than that of others. The opportunity given each child to be educated should bear a relation to his educability. Thus some should be kept at school longer than others; some should be given secondary school education while others should not.

In order to study the question of a child being educated in proportion to his relative mental ability, the destination of all tested pupils who had completed the Public School Course was ascertained, as far as possible. It was possible to secure definite, exact information regarding 75 such pupils. Of these, 58 had gone to some institution giving secondary education. The remaining 17 went to work after completing the Public School Course. The average I. Q. of those who attended a secondary school was 103.6; their range of I. Q. was from 79 to 137. The average for those who went to work was 97.3; their range of I. Q. was from 82 to 127. Only 15.5% of those who kept on going to school had I. Q.'s below 90, while 41.2% of those who went to work had I. Q.'s below 90. Thus the tendency is for the brighter pupils to get an education beyond the public schools, and for the duller pupils to go to work. One notices a few exceptions to this tendency. The boy with an I. Q. of 79 went on to high school while the boy with an I. Q. of 127 went to work as messenger boy for a butcher. There surely should be provision made, by means of scholarships, special grants, etc., whereby a deserving child with con-

siderable ability might go on to school. At the present time, after working for the butcher for two years, the boy referred to is earning nine dollars a week. A small committee of educationists or citizens with a little money to spend, or some of the numerous clubs and societies whose purpose is to help those who need help, could do a great work by making a secondary school education possible for poor, deserving, clever children.

The school whose pupils were tested has an average enrolment of 725 pupils. The testing was voluntary and was done after school hours and on Saturday mornings; it extended over a four-year period. In many cases members of the same family were examined; all the pupils, especially the seniors, heard about the testing from time to time, and doubtless talked a great deal about it. One can imagine that the questions and possible answers would be discussed by the pupils. Did coaching and discussion of the tests by the pupils have any effect on the results? The examiner watched carefully for any evidences of coaching. In only one test was there reason to believe that some had been helped. Some pupils passed the "Ball and Field" test of Year XII in such a way as to arouse the suspicion that they had been told of it previously. To get definite facts on the matter, the average I. Q. of the first class tested was ascertained and also the average of the last. In both cases the number of pupils in the class was 35. The first class tested averaged 99.4 and the last one averaged 98.4. If there had been any improvement resulting from discussing the tests, the last class would certainly have benefited by it; yet their average I. Q. was lower than that of the first class tested. Thus it is evident that the effect of the children's discussing the tests was negligible.

All the children in a school should be tested, if the results are to be used in school work. What kind of test should be used? If the number of pupils be small, the individual Terman test should prove very useful. In a large school this is not always possible because the time a teacher can spend at testing is necessarily limited. The average time used to test those reported was about forty-five minutes for each pupil. The shortest time recorded was twenty-four minutes, and the longest an hour and a half. Thus to test 500 pupils required about 375 hours; another 25 hours—an average of three minutes per pupil—was needed to score the answers and calculate the I. Q.'s Thus the time required limits the use of such individual tests as Terman's. As previously pointed out, the Otis Group Test is highly accurate, and requires but a short time to administer to a group. More time, however, is needed to score the Otis test papers than the Terman. But allowing for this, the time required is not such as to be prohibitive. Five hundred pupils could be given the Otis tests and the answers scored in about fifty hours. Thus in a large school the best method is to give a group test to all pupils. Those who are ranked very high or very low should then be given an individual test and the results compared. The relation between the results of the two tests should, of course, be known; for example, the Otis I. Q. is about 10% higher than the Terman I. Q. All pupils should be given two or three group tests at intervals of about a year. If the results of three tests are approximately equal for a child, the average of the three may be accepted as accurate. Thus with a comparatively small expenditure of time each year, accurate test results for a large number may be determined.

One who uses them realizes that mental tests are not infallible. But the knowledge gained by their use amply repays one for the time spent. Every child tested impresses on the examiner that each one is mentally different from every other one. Tests that an average child will pass quite easily will often be difficult for a bright child. A dull child will often correctly answer a fairly difficult test. One is amazed at the variety of mental abilities shown by various pupils. The complexity of the whole subject of intelligence is constantly emphasized. One often finds children who appear very inconsistent. For example, a boy of sixteen, when tested, had an I. Q. of 69. The calibre of his school work stamped him as being very dull, if not mentally defective. Yet how he could play the piano! His ability with the piano was much greater than that of many older people who had had more musical training than he had had. A girl with an I. Q. of 97 and a mental age of 13 years 10 months led her class in monthly examinations throughout the year. Several in that class had I. Q.'s over 100. Several were mentally three years older than that particular girl. Other brighter pupils worked hard but could not rank with her on the examinations. The reason likely was that she worked very hard, and also that she was a stolid type; she never became excited or "lost her head", and this quality was an advantage during examinations. The degree of application and the emotional condition of a child are two of the most important factors in his school progress.

A teacher may think he knows and understands his pupils very well before testing them. He will know

them much better after testing them. By an intelligence test an examiner will frequently get a better understanding of a child's mental life and his mental ability than he would have by teaching him for six months. The better a teacher understands a child's mental capacities, the better he should be able to teach him. From the standpoint of the increased knowledge of a child which is acquired by the examiner during a test, intelligence testing is well worth the time spent on it. One often wishes he could go further than test. He may find a child lacking in memory ability, or in power to associate, or concentrate, or to imagine, or compare concepts, etc. The next logical step would be to prescribe a remedy for such deficiencies. A book by a competent authority designed for teachers' use, and dealing with methods of overcoming various mental handicaps, would be very valuable. A teacher could then be a sort of mental physician. He could diagnose the mental condition of the child and then prescribe the necessary steps to overcome any apparent deficiencies.

Some children do not make scores which are as high as they possibly could make with their respective native abilities. One reason for this is that responses are often indefinite. The examiner is, in most tests, permitted to question the pupil to some extent so as to ascertain his real meaning: however, there is a limit to the amount of questioning an examiner can do; generally speaking, he must strictly avoid any question or remark which would indicate to the child that his answers were either right or wrong. To answer many questions correctly, a thoughtful, definite answer from the child is necessary; often a child

fails to score in a test because his answer does not accurately express the thought he has. It would seem advisable for teachers to devote more time and attention to oral composition, and to insist on definite, lucid answers from pupils. Children should be more thoroughly trained in careful, correct expression of thought.

The other chief defects noticed as general among those examined were a deficiency in general information, a poor vocabulary, and a lack of ability to concentrate. Many teachers complain that this last defect is much more noticeable than it used to be among children. Most teachers blame the moving picture shows for it. Moving pictures often train a person not to concentrate. One scene is flashed on the screen after another and one needs to use little, if any, mental effort to follow most stories. Many children visit moving picture theatres two or three times each week, and the frequency of such visits, along with the training they receive while there, does not encourage the exercise or development of desirable mental qualities.

The three defects noted in the last paragraph could be overcome to some extent by giving more attention to the reading done by a child. A larger vocabulary and more general information could be acquired by more reading. The power of concentrating could be developed by training a child to remember what he has read. Writing a summary of a page or a chapter of his reading would be a valuable exercise as it would cause the child to think over what he has read. This exercise would prove very valuable in developing several of a child's mental capacities.

The place of intelligence tests in education has been

established. In the Introduction reference was made to the wonderful increase in their use during the last few years. Mr. H. C. Link, writing in the *Atlantic Monthly*, September, 1923, states: "In view of these achievements, I have no hesitation in expressing the opinion that Professor Terman and the psychologists responsible for the development of intelligence tests, have made the greatest single contribution to the field of education in our time."

Is there work in this field for Canadian teachers and educators? Most assuredly there is. We need Canadian intelligence tests. Some of the Terman tests are not well adapted for use with Canadian children. An example of such a test occurs in those for year 14: the child is asked to name three differences between a president and a king; the correct answer must include at least two of the differences relating to power, accessions, or tenure. Very few of those tested gave correct answers. Such information may be within the range of experience of the average fourteen-year-old child in California, but such is not the case when applied to the children of Canada. Canadian children are quite ignorant about the tenure or accession of presidents, but most of those tested readily expressed the conviction that a king is much more powerful than a president; other differences mentioned by many were that a king wore a crown and lived in a palace, while a president did neither of these things.

The range of experience of the children of one country differs somewhat from the experiences common to the children of another country. This explains the necessity for Canadian intelligence tests. If we are to have them, Canadian teachers must provide them. Just as the province of Ontario has led in many Cana-

dian educational movements in the past, so should Ontario lead in this new field. The teachers of Ontario are earnest and zealous and will nobly do their part in any movement which will improve the work so dear to them. To devise a system of intelligence tests needs the co-operation of dozens, or better, hundreds, of teachers. Leadership is needed. For this, one looks either to some zealous member, or group of members, of the teaching profession; or to some central body, such as a Normal School or the Provincial University. May Canadian teachers soon seriously take up the study and use of intelligence tests! Then Canadian schools will experience the benefits that have followed their use in other school systems.

BIBLIOGRAPHY

Anderson: Estimating Intelligence by Means of Printed Photographs: *Journal of Applied Psychology*, June, 1921.

Ballard: Mental Tests, Hodder & Stoughton, London, 1920.

Ballard: Group Tests of Intelligence, Hodder & Stoughton, London, 1922.

Franzen: The Accomplishment Ratio, Teachers' College, New York, 1922.

Galton: Inquiries into Human Faculty, J. M. Dent, London, 1883.

Garrison and Tippett: Comparison of the Binet-Simon and Otis Tests; *Journal of Educational Research*, June, 1922.

Journal of Educational Psychology, Vol. XII: Intelligence and Its Measurement: A Symposium.

Memoirs: National Academy of Sciences, Washington, 1921.

Murdock: The Accomplishment Quotient; *Teachers' College Record*, May, 1922.

Pintner: Intelligence Testing, Holt & Co., New York, 1923.

Seventeenth Year Book of the National Society for the Study of Education; The Measurement of Educational Products.

Starch: Educational Psychology, Macmillan Co., New York, 1920.

Stebbins and Pechstein: Quotients, I. E. and A.; *Journal of Educational Psychology*, October, 1922.

Symonds: Intelligence of Population in United States; *Journal of Educational Psychology*, February, 1923.

Terman: The Measurement of Intelligence, Riverside Press, Cambridge, Mass., 1916.

Terman *et al* Intelligence Tests and School Reorganization, World Book Co., Yonkers-on-Hudson, 1923.

Toops and Symonds: What Shall We Expect of the A. Q.? *Journal of Educational Psychology*, December, 1922; January, 1923.

Whipple: Manual of Mental and Physical Tests, Warwick and York, 1914-15.

Woodrow: Brightness and Dullness in Children, Lippincott Co., Philadelphia, 1919.

Wylie: A Brief History of Mental Tests; *Teachers' College Record*, January, 1922.

Yoakum and Yerkes: Army Mental Tests, Holt & Co., New York, 1920.

ioned

n a Toronto Public School.